THE WORLD BANK ECONOMIC REVIEW

Volume 16 • 2002 • Number 3

www.wber.oupjournals.org

OXFORD

ISSN 0258-6770

THE WORLD BANK ECONOMIC REVIEW

The World Bank Economic Review is a professional journal for the dissemination of World Bank–sponsored and outside research that may inform policy analyses and choices It is directed to an international readership among economists and social scientists in government, business, and international agencies, as well as in universities and development research institutions The *Review* emphasizes policy relevance and operational aspects of economics, rather than primarily theoretical and methodological issues It is intended for readers familiar with economic theory and analysis but not necessarily proficient in advanced mathematical or econometric techniques Articles will illustrate how professional research can shed light on policy choices Inconsistency with Bank policy will not be grounds for rejection of an article

Articles will be drawn from work conducted by World Bank staff and consultants and from papers submitted by outside researchers Before being accepted for publication, all articles will be reviewed by two referees who are not members of the Bank's staff and one World Bank staff member Articles must also be recommended by a member of the Editorial Board Non-Bank contributors are requested to submit a proposal of not more than two pages in length to the Editor or a member of the Editorial Board before sending in their paper

Comments or brief notes responding to *Review* articles are welcome and will be considered for publication to the extent that space permits Please direct all editorial correspondence to the Editor, *The World Bank Economic Review*, The World Bank, 1818 H Street, Washington, DC 20433, USA, or wber@worldbank org

---

For more information, please visit the Web sites of the *Economic Review* at www.wber.oupjournals org, the World Bank at www worldbank org, and Oxford University Press at www oup-usa org

THE WORLD BANK ECONOMIC REVIEW

Volume 16 • 2002 • Number 3

The World Bank Economic Review (ISSN 0258-6770) is published three times a year by Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009 for The International Bank for Reconstruction and Development / THE WORLD BANK. Communications regarding original articles and editorial management should be addressed to The Editor, *The World Bank Economic Review*, 66, avenue d'Iéna, 75116 Paris, France. E-mail: wber@worldbank.org.

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide.

SUBSCRIPTIONS: Subscription is on a yearly basis. The annual rates are US$40 (£30 in UK and Europe) for individuals; US$101 (£71 in UK and Europe) for academic libraries; US$123 (£84 in UK and Europe) for corporations. Single issues are available for US$16 (£12 in UK and Europe) for individuals; US$43 (£30 in UK and Europe) for academic libraries; US$52 (£35 in UK and Europe) for corporations All prices include postage. Individual rates are applicable only when a subscription is for individual use and are not available if delivery is made to a corporate address. *Subscriptions are provided free of charge to non-OECD countries.* All subscription requests, single issue orders, changes of address, and claims for missing issues should be sent to:

North America· Oxford University Press, Journals Customer Service, 2001 Evans Road, Cary, NC 27513-2009, USA. Toll-free in the USA and Canada: 800-852-7323, or 919-677-0977. Fax: 919-677-1714. E-mail: jnlorders@oup-usa.org.

Elsewhere Oxford University Press, Journals Subscriptions Department, Great Clarendon Street, Oxford OX2 6DP, UK. Tel· +44 1865 353907. Fax: +44 1865 353485. E-mail: jnls.cust.serv@oup co.uk.

ADVERTISING. Helen Pearson, Oxford Journals Advertising, P.O Box 347, Abingdon SO, OX14 1GJ, UK Tel/Fax: +44 1235 201904. E-mail· helen@oxfordads.com.

BACK ISSUES: The current plus all back volumes (from 1997) are available from Oxford University Press at the *North America* contact information listed above.

INDEXING AND ABSTRACTING: *The World Bank Economic Review* is indexed and/or abstracted by *CAB Abstracts*, *Current Contents/Social and Behavioral Sciences*, *Journal of Economic Literature/EconLit*, *PAIS International*, RePEc (*Research in Economic Papers*), and *Social Sciences Citation Index*. The microform edition is available through ProQuest (formerly UMI), 300 North Zeeb Road, Ann Arbor, MI 48106, USA.

PAPER USED: *The World Bank Economic Review* is printed on acid-free paper that meets the minimum requirements of ANSI Standard Z39.48-1984 (Permanence of Paper).

POSTAL INFORMATION: *The World Bank Economic Review* (ISSN 0258-6770) is published three times a year by Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009. Send address changes to *The World Bank Economic Review*, Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009.

THE WORLD BANK ECONOMIC REVIEW, VOL. 16, NO. 3 321–343

Gender Effects of Social Security Reform in Chile

Alejandra Cox Edwards

In 1981 Chile replaced a mature government-run social security system that operated on a pay-as-you-go basis with a privately managed system based on individual retirement accounts. The new system is more fiscally sustainable because pension benefits are defined by contributions. The minimum pension guaranteed to beneficiaries with at least 20 years is funded from general taxes, preserving the tight matching between contributions and benefits. The new system also eliminates several cross-subsidies. Men and women with less than secondary education gain under the new system, but single women with more education lose. Comparison of the old and the new systems reveals a complex set of factors that cause gender effects given constant behavior or change behavior across genders.

I. Chile's Social Security System

By the late 1970s Chile's social security system had generated a large deficit despite several hikes in payroll taxes. In 1979 pension ages were raised to put the system in balance (Wagner 1983). In 1981 Chile replaced the mature government-run social security system, which operated on a pay-as-you-go basis, with a privately managed system based on individual retirement accounts. The new system is more fiscally sustainable than the old system, with contribution-defined pension benefits. The reformed system includes a minimum pension guarantee, which is funded from the general government budget, preserving the tight link between contributions and benefits of the pension funds. The reformed system also insures contributors against the risk of loosing their income-generating capacity before retirement age by including an explicit insurance premium covering the risks of disability and early death.

From the point of view of macroeconomic aggregates, the Chilean social security system—managed mostly by private companies, Administradoras de

Alejandra Cox Edwards is in the Department of Economics at California State University, Long Beach Her e-mail address is edwards.ac@attbi.com. This article draws heavily on two reports prepared for Gender and Social Security, a World Bank research project managed by Estelle James (World Bank Institute) and Maria Correia (Latin America and the Caribbean Region Gender Team) The author is grateful to Estelle James, Salvador Valdes, anonymous referees, and seminar participants at Centro de Estudios Públicos, Santiago, Chile, for their comments and suggestions and to the Economics Department at Universidad de Chile, particularly Dante Contreras, for assistance with the CASEN data.

DOI 10.1093/wber/lhf002

Fondos de Pensiones—can have annual surpluses or deficits. The long-term viability of the system is protected by the direct link between contributions and benefits. From the individual's perspective, social security contributions have become a form of delayed compensation because all parts of the contribution have a counterpart in paid benefits.[1] This change in design represents a reduction in the payroll tax. In addition, the reform reduced the payroll tax rate earmarked to finance pensions and health insurance from about 30 percent to 20 percent of taxable earnings.[2]

Social security systems have particular ways of dealing with intrafamily distribution of incomes and old-age benefits. To address these differences, this study estimates individual contributions and benefits before and after reform. The focus is on gender differences in contributions and old-age benefits; other social security benefits (such as disability and premature death) are left out of the comparison. I examine the effects of four changes:

1. Workers who contributed for less than 10 years received no pensions under the old system. Under the new system everyone who has contributed to the system receives a pension, in proportion to his or her fund accumulation. This change is particularly significant for women, who tend to have low levels of attachment to the labor force.
2. In the old system, pension benefits were determined by a formula that multiplied the average contributory income of the last five years of work by the number of years of contributions, rewarding a long-term commitment to the labor market. In the new system, pensions are a function of the accumulation of funds through compound interest, giving a heavier weight to contributions made early in life. This change favors women as well as men who have career interruptions or relatively flat age-earnings profiles throughout their careers.
3. Unlike the old system, in which survivor's pensions were funded from systemwide contributions, the new system is based on joint annuities. That is, the pension benefits of the principal are based on the pensioner's fund accumulation, net of a reserve to fund his or her survivor's benefits. This change internalizes the accumulation of pension funds at the family level,

1. The system is still perceived as a tax by people whose desired savings rate is below the implicit rate imposed by social security. Edwards and Edwards (2002) estimate the tax component of the Chilean payroll contribution toward social security.

2. The old system was financed by a payroll tax, with a relatively weak link between contributions and benefits. In 1973, for example, contributions to the retirement plan by employers and employees averaged 26 percent of earnings. Once contributions to the national health system were included, total payroll contributions exceeded 50 percent of earnings for some workers. During the late 1970s payroll tax rates were lowered; in 1980 social security contributions claimed 32.50–41.04 percent of taxable earnings. The new system reduced the overall contribution to social security (pension, health, and other forms of insurance) to about 20 percent of taxable earnings, established a set of common rules for all contributors, compartmentalized the various parts of the social security package in different products, and introduced competitive forces in the market for these products.

increasing the marginal benefit to the family of every peso put into the system.

4. In the old system's main program, widows were required to choose between survivor's benefits and their own retirement benefits. This rule reduced the marginal value of their own contributions, making the payroll tax even higher for married women. Under the new system, benefits can be combined, improving work incentives for married women.

These four changes have raised the marginal benefit of own contributions for all, particularly for those whose contributions were insufficient to qualify for benefits. Furthermore, the new system guarantees a minimum pension to people who contribute at least 20 years and whose fund accumulation provides for an annuity below the minimum pension. This represents an additional incentive for participation, although it also acts as a disincentive to participate and contribute once a person with relatively low contributions has accumulated 20 years in the system.

The current living standards of the elderly population are the product of both a traditional system of extended family arrangements and a social security system that proved to be unsustainable and was therefore fully reformed. In urban areas the proportion of elderly men who receive an old-age pensions (62 percent) is twice that of elderly women (31 percent). Another 19 percent of elderly women receive survivor's pensions, closing the retirement income gender gap. In rural areas, where old-age pensions are less typical, more than 23 percent of elderly women are beneficiaries of pensión asistencial (PASIS), a government program targeting the elderly poor that operates outside of the social security system funding. In urban and rural areas older women are more likely than men of the same age group to be widowed and to live in extended households (Edwards 2000).

Future living standards for elderly women will depend on living arrangements, the coverage of the PASIS program or its equivalent, and the pension system. Future pension benefits for women are determined by two main sources: their own accumulated funds and survivor benefits for married women. As schooling levels have risen in Chile, marriage age has been delayed and labor force participation rates have increased. The fraction of women qualifying for pension benefits is therefore likely to increase.

Women who work for pay typically accumulate pension funds at a significantly lower pace than men for two reasons. First, women are more likely than men to interrupt their careers (to take care of children, the sick, or elderly parents or in-laws). As a result, the density of labor market participation throughout the life cycle is lower for women than for men. Second, women's salaries are generally lower than men's, even after controlling for age and schooling. The already slower accumulation is often stopped at a younger age, given that the social security system allows women to collect pensions at age 60, whereas men must wait until age 65. At age 60 women are expected to live 23 more years, a long stretch for the accumulated funds. However, the present value of expected

old-age benefits of a married woman is above the level of accumulated funds because of the likelihood that she will survive her husband, whose pension she will receive.

II. Data and Methodology

The key data source for the analysis is the micro data set of the Caracterización Socioeconómica Nacional (CASEN) for 1994, a national household survey carried by the National Planning Office. This survey collects information on a variety of indicators, including demographic characteristics, labor force participation, earnings, affiliation to social security, and the answer to the question "Are you a contributor to any of the social security systems?"

CASEN is a reliable source for estimating the socioeconomic characteristics of the population. The other source of information on affiliates and contributors is the Superintendencia de Administradoras de Fondos de Pensiones (SAFP), to which all private Administradoras de Fondos de Pensiones provide information on their accounts. The SAFP publishes aggregate data and does not provide researchers with individual-level data. According to SAFP data, the ratio of contributors to affiliates fell from 58 percent in 1985 to 49 percent in 1994 and to 44 percent in 2000. This decline is driven by the fact that affiliation is a forever classification. People who join the labor force and make contributions for a short time and then do not return to the labor force remain affiliated with the system, even though they may forget that they belong to it. Even if people leave the country, they remain in the SAFP counts.[3]

Unfortunately, information on years of contributions is not available from CASEN or publicly available from the SAFPs. This article uses the cross-section data to build a series of synthetic cohorts and use them to project life-cycle earnings and contributions of "typical" people. The methodology consists of defining key observed individual characteristics and measuring employment patterns by age, leading to the construction of synthetic cohorts from the cross-section.

The cross-section data allow us to estimate the labor force participation rate for a typical woman of a given age. The observed participation patterns can be used to project the behavior of a 20-year-old into the future, assuming that the behavior of the observed 30-, 40-, and 50-year-olds is characteristic of all women. Young women today are unlikely to behave as their predecessors did, however, particularly regarding labor force participation. The key factor driving this generational change is that the younger cohorts have more schooling. Thus unless one controls for schooling, a synthetic cohort built from cross-section data would introduce an error in the link between age and labor force participation.

3. As an anonymous referee pointed out, the proportion of contributors to affiliates based on CASEN data is 55 percent for the national data—6 points above the SAFP estimate. I do not believe this difference invalidates the methodology. Although CASEN excludes affiliates who have left the country or forgotten their affiliation, its estimate is still representative of typical working affiliates.

After carefully studying the links between labor force participation and schooling, I divided the 1994 urban sample by gender and into five schooling categories: incomplete primary, incomplete secondary, complete secondary, up to four years of postsecondary, and more than four years of postsecondary. The assumption is that each of these schooling groups has a common pattern of labor force participation and that the pattern is stable over time. If the average rate of employment over a period is used to estimate the number of months of work, the implicit assumption is that all people in the sample base work some of the time. This assumption is appropriate for the men's cases, because it is known (from cohort data) that practically all men have been in the labor force for some time by age 30 (Edwards 2001). However, evidence from recent cohorts suggests that no more than 95 percent of women with postsecondary schooling join the labor force at some point, and this fraction falls to 80 percent and 75 percent for females with less schooling. Therefore, this procedure would underestimate the number of months of work for the "appropriate" sample of women.

Information on affiliation is very useful because affiliated men and women are known to have been in the labor force at some point. This variable allows us to separate men and women into two subgroups: people who are affiliated with the system and make contributions toward pensions at least some of the time and the unaffiliated. Naturally, participation rates are higher among women affiliated with the social security system relative to all women. Given our goal of estimating social security contributions and benefits, the methodology rests on estimates for women affiliated with the social security system by schooling and marital status.

Final values of accumulated contributions are a function of the system's rate of return and individual retirement ages. Individual pension benefits estimated from the accumulated funds depend on survival probabilities at pension age. Estimated individual benefits are subsequently compared with what the same "typical" people would have obtained using the nominal formula of the old system. These comparisons allow us to better understand the gender effects of social security reform.

III. Employment and Social Security Participation by Gender

Labor force participants are a subsample of working-age people with distinct gender, age, schooling, and marital status characteristics. Labor force participation of 16- to 65-year-old urban women and men is driven by marital status, years since finishing school or potential experience, and postsecondary schooling. Participation is significantly lower for women (39 percent) than for men (82 percent). Marriage further reduces the probability of participation for women and increases it for men. Postsecondary schooling increases the likelihood of participation and diminishes the negative effect of marriage for women but lowers the probability of participation for men.

Who Is Required to Affiliate and Who Contributes to Social Security?

Chilean law requires formal employees to make contributions to their retirement accounts, and allows the self-employed to make voluntary contributions to the pension system.[4] Affiliation is necessary to contribute to the system and to obtain benefits, and once a person affiliates to the system, he or she remains affiliated for life. In 1994, 67 percent of men and 39 percent of women in the working-age population were affiliated with the system. Affiliates have a higher than average attachment to the labor force. About 82 percent of male affiliates and 64 percent of women affiliates were working at the time of the survey—higher percentages than the 72 percent of men and 36 percent of women working in the total working-age population.

Unlike social security affiliation, work status is not set for life. Many people become self-employed after working as employees or move back and forth between the two categories. It is therefore not surprising that a significant fraction of workers classified as self-employed make contributions to social security. In 1994 about 25 percent of people who were not required to contribute (the self-employed and employees without contracts) did so. The fraction of self-employed people who make contributions increases with age and schooling. Contributing does not vary much with the level of salary or the sector of employment, but establishment size and gender are important factors. In particular, workers in larger establishments are more likely to contribute. With or without controlling for industry and establishment size, self-employed women are 6.2 percent less likely to contribute than self-employed males.[5]

This study estimates working patterns of people affiliated with the social security system and assume that month-to-month contributions by affiliates are driven primarily by whether they are working in a given month. In fact we know that more than 90 percent of men and women affiliates who were working at the time of the survey contribute toward social security. There is a remarkable similarity across genders in contributory status, as long as the sample is limited to affiliated individuals. Among men and women with less than primary schooling, 84 percent contributed in 1994. Contribution rates rose to 89 percent among those with incomplete primary, to 93 percent among those with complete secondary and men with up to four years of postsecondary education, and to 96 percent for women with postsecondary and men with five or more years of postsecondary education.

4. Formality is established by a written contract that employers and employees are required to sign.

5. This finding suggests that in similar circumstances to men, women are less likely to assign a marginal value to their social security contributions. There are two possible explanations First, there may be a relatively larger fraction of women who work for pay in a given period who are not planning to work for pay for any significant length of time. The fact that married women obtain health care coverage through a contributing husband's family plan may also be part of the explanation. This reduces the value of the 20 percent contribution to about 13 percent. Further research should examine the impact of the tied-in character of the pension and health care programs on a married couple's incentive to save toward retirement.

Estimating Years of Contributions

The probability of employment varies by schooling, marital status, and gender. The sample of affiliates is divided into 20 main categories based on these three variables. Within each of these categories, I calculate the fraction of men and women that work at every age. I then assume that the typical man or woman within each schooling category works as a typical single person until marriage and as a married person afterward. The marriage age for the typical man and woman is defined as the age at which 50 percent of the corresponding category is married. Marriage tends to be earlier for men affiliates than for men as a whole, and marriage tends to be delayed for women affiliates compared with the women as a whole. Because I am focused on affiliates, I use the corresponding marriage ages for that sample. This step reduces the 20 categories to 10, 5 schooling categories and 2 genders. The average fraction of workers is used to estimate the "fraction of the time that typical individuals work at every age." The results of this estimation are summarized in table 1 and expressed in estimated years of work by age category.

Assuming that working affiliates make contributions to social security, the emerging relation between age and years of contributions is captured in figure 1. A higher line in figure 1 indicates steady accumulation of contributions over many years, as is the case for men with secondary schooling. A relatively high and steep line indicates rapid accumulation after a later start, as is the case for men and women with five or more years of postsecondary schooling. A lower and flatter line shows a lower degree of labor force attachment and a relatively slower accumulation of contributions, as is the case with women with secondary schooling.

Gender differences in the degree of attachment to the labor force among affiliates result in important differences in estimated lifetime contributions. Men typically accumulate 40 years' worth of contributions between age 16 and 65. Women, especially women in the lower schooling categories, tend to have more interruptions. As a result, on average women who complete secondary schooling accumulate less than 30 years' worth of contributions by age 65.

Earnings Profiles

To estimate the accumulation of funds, I use the estimates of contributory behavior—obtained in the previous subsection—and earnings. I produce an estimate of monthly wages by age, based on observed earnings by age, sex, and schooling. The sample includes all workers in an attempt to keep it as large as possible. This is appropriate because wage levels do not significantly affect affiliation and contribution behavior. Thus earnings estimates originating from a broad sample of workers should not differ from estimates originating from a sample of ever working affiliates.

I start from the assumption that current patterns of earnings (as a function of schooling and experience) have persisted for some time and will remain stable in the future. The key challenge is to capture the earnings pattern from the exist-

TABLE 1. Estimated Years of Contributions, by Age and Schooling

Age category	Incomplete primary	Incomplete secondary	Complete secondary	Up to 4 years postsecondary	More than 4 years postsecondary
Ever working male affiliates					
16–20	3.37	3.02	1.65	1.28	0.00
21–25	7.02	6.91	5.84	5.35	2.79
26–30	10.65	11.04	10.33	9.77	7.25
31–35	14.49	15.33	15.03	14.47	12.12
36–40	18.46	19.63	19.72	19.09	17.02
41–45	22.86	23.98	24.25	23.56	21.77
46–50	26.75	28.22	28.37	27.79	26.48
51–55	30.49	32.42	32.34	31.80	31.25
56–60	33.52	35.74	36.04	35.63	35.91
61–65	35.98	38.05	38.29	38.97	38.97
Ever working female affiliates					
16–20	3.64	2.85	1.39	1 21	0.00
21–25	6.84	6.51	5.31	5.12	3.26
26–30	9.80	10.03	8.64	8.90	7.87
31–35	11.44	11.95	11.58	12.64	12.10
36–40	13.79	14.78	14.95	16.35	16.92
41–45	16.23	17.69	18.32	20.72	21.53
46–50	18.32	19.90	21.70	24.08	26.18
51–55	20.86	22.21	24.17	28.52	30.81
56–60	22.49	24.06	26.55	32.68	34.77
61–65	23.42	24 17	26.80	32.92	36.05

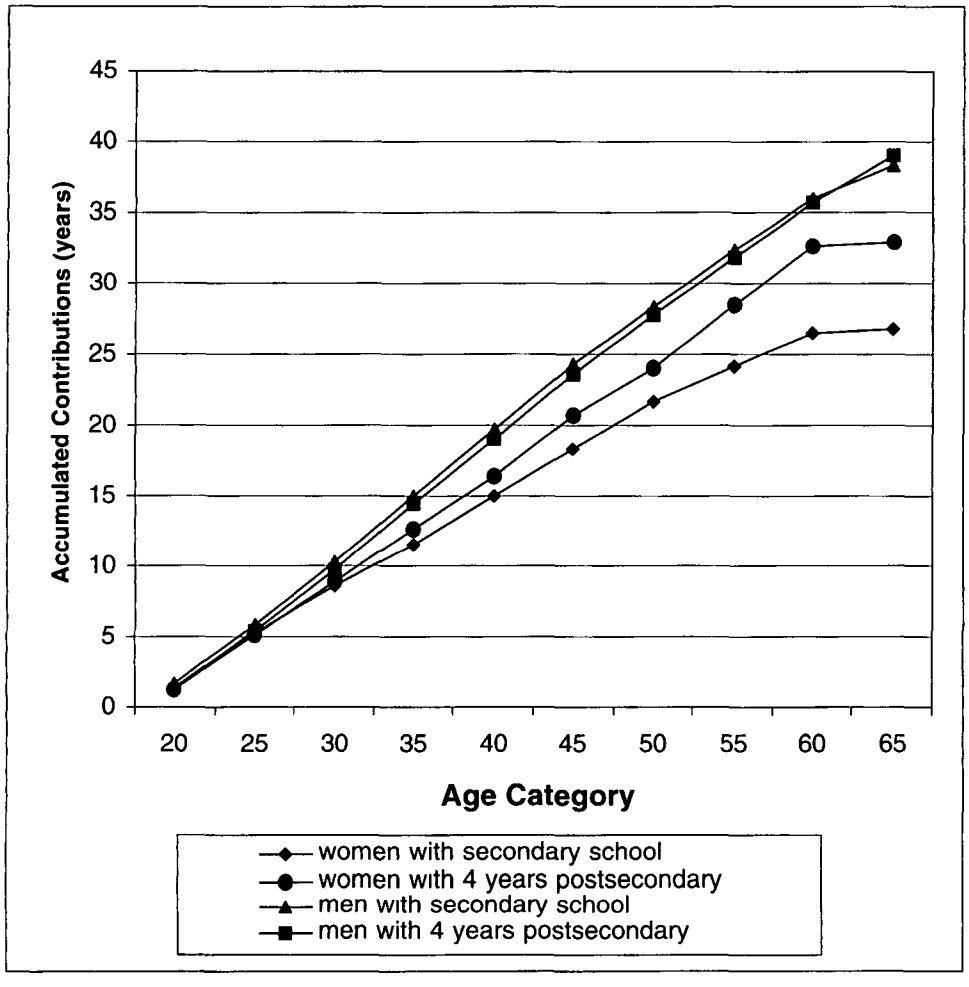
Note: Data are for urban areas only.
Source: Author's estimates based on CASEN 94 data.

ing data. The human capital earnings function, in which earnings are expressed as a quadratic in potential experience, is probably the most widely accepted empirical specification in economics. This procedure is not the most appropriate here, however, because I lack a good proxy for female experience and my aim is to get the best estimate of earnings for workers of a given age (because contributions and benefits eligibility are bound by age). I therefore use an alternative procedure to compute earnings.

I organize the data on earnings from the 1994 CASEN survey by sex, age, and schooling and calculate an average income for each cell (table 2). This method does not impose a particular functional form, and it has the advantage of implicitly weighting the sample according to its composition (by other characteristics) within each cell. Given the limitation imposed by sample sizes, it is not possible to estimate average earnings for single-age categories. A five-year interval was chosen to increase sample size while keeping the age categories narrow, because estimated salaries for a range of years are likely to overestimate starting-period contributions and underestimate end-period contributions.

The resulting earnings-experience profiles have a concave shape: Earnings grow fastest at the earlier stage of most groups' careers, earnings growth slows after age 40, and earnings often fall bellow the peak by age 60. Estimated gender in-

FIGURE 1. Estimated Years of Contribution by Age, Schooling, and Gender



Source: Table 1.

come differentials by schooling and age indicate that the female-to-male income ratio is 0.6–0.8 in most cases. The notable exception is people with five or more years of schooling, among whom the differential is closer to 0.5. The income differential grows sharply between ages 45 and 50 and then declines again, a feature that affects savings accumulation toward pensions.

IV. Fund Accumulation and the Effect of Gender on Pensions

This section explores the effect of gender on the accumulation of pension funds, pension benefits, and replacement rates, based on simulations for representa-

TABLE 2. Average Income, by Age and Schooling
(1994 pesos, except where otherwise indicated)

Age category	Incomplete primary	Incomplete secondary	Complete secondary	Up to 4 years postsecondary	More than 4 years postsecondary
Estimated monthly male earnings					
16–20	49,145	61,958	72,894	76,149	n.a.
21–25	61,366	72,884	89,050	119,020	313,293
26–30	67,259	84,219	108,092	155,493	358,164
31–35	70,030	94,988	133,436	195,497	482,094
36–40	76,019	103,699	151,606	223,750	524,083
41–45	88,323	115,844	174,791	248,305	540,316
46–50	93,893	143,450	221,171	269,793	643,224
51–55	90,986	128,078	201,733	247,731	595,663
56–60	92,653	135,883	197,906	281,721	542,736
61–65	81,430	122,726	161,457	240,541	513,568
Estimated monthly female earnings					
16–20	48,479	48,124	62,718	66,393	n.a.
21–25	49,496	60,800	75,702	95,447	179,198
26–30	53,374	59,136	82,812	167,499	232,048
31–35	53,044	66,317	91,003	130,258	260,202
36–40	52,251	70,051	107,584	138,252	304,915
41–45	58,110	79,232	137,248	179,873	312,696
46–50	60,745	83,353	134,975	209,027	212,333
51–55	62,959	75,782	156,673	154,783	222,027
56–60	63,795	93,730	168,694	149,990	283,680
61–65	58,703	62,813	116,958	157,500	365,000

n.a. = Not applicable.
Note: Data are for urban areas only and are based on full-time earners.
Source: Author's estimates based on CASEN 94 data.

tive workers. The impact of the change from a pay-as-you-go to a multipillar system and of particular pay-out policies of a defined-contributions system is calculated.

Gender Differences in Fund Accumulation

I assume that workers in a given schooling and gender category contribute 10 percent of their income, as required by law. Earnings for each age are assumed to be equal to the estimated value for the corresponding five-year age period. In some simulations I add a secular growth to earnings, increasing the estimated annual wage by the corresponding growth effect. For a given age, estimated annual months of contributions are equal to 12 times accumulated contributions for the five-year period divided by 5. The accumulation of funds is the result of compounding the estimated contributions at various interest rates.

According to these estimates, women accumulate funds at a lower pace and have income profiles that are flatter and lower than those of men. The estimates

presented are to be used as a benchmark; the system's rates of return as well as the affiliate's income and number of years of accumulation relative to the mean affect these figures. Given the same estimated earnings, the higher the rate of return, the larger the accumulated fund. The longer a person works, the higher the annuity, with benefits falling more than proportionally if a person works less than the 20 years required to qualify for the minimum pension. (For a broader set of estimates, see Edwards 2001.)

The calculations in table 3 show estimated lifetime accumulation of funds for men and women by level of schooling. Estimates based on a 5 percent rate of return and a 2 percent income growth generate women's accumulations that are 36–52 percent of men's. At the bottom of the table, I decompose the difference in fund accumulations within each schooling category in four steps:

1. If women postpone retirement to age 65, the gap between men and women would narrow by 10–15 percent depending on the level of schooling. This effect tends to be larger for more educated women because their work intensity from age 60 to 65 tends to be higher.
2. If women continue to work only to age 60 but do so with the same work intensity of men, their pension funds would increase 1–19 percent, depending on the level of schooling. This effect varies significantly across schooling groups, because there is a significant variation in work intensity among women by schooling group, with highly educated women working almost as intensively as men with the same level of education.
3. If women work to age 60 and do not change their work patterns but are paid the same as men, the pension fund gap between men and women would narrow by 7–25 percentage points, depending on the schooling group. The gap would narrow 7 percent among people with secondary education and up to four years of postsecondary education, by about 13 percent for the lower education groups, and by more than 25 percent for people with the highest level of schooling. One reason for the larger effect of income increases for women in the highest schooling group is that this is the group with the most labor market attachment. An increase in income levels is thus weighted by a higher number than an increase in earnings of other groups of women.
4. The last step calculates the effects of the interaction of the first three effects.

Estimated Future Pension Benefits

The combination of earlier retirement age and longer life expectancy means that women need to provide for 23 years of income from their accumulated funds. Married men retire at age 65 and must provide for a joint annuity of 21 years. This joint annuity is composed of 15 years of own pension and 6 years of widow's pension, at 60 percent of own benefit. Single men retire at age 65 and must provide for their own 15 years' annuity. Therefore, even if men and women started with the same fund accumulation at their "normal" retirement age, a woman's

TABLE 3. Gender Differences in Fund Accumulation
(thousand of 1994 pesos, at 413.45 pesos per US$)

Item	Incomplete primary	Incomplete secondary	Secondary	Up to 4 years postsecondary	More than 4 years postsecondary
Women retire at 60 and retain female work patterns (pesos)	5,612	7,832	13,808	22,801	41,918
Women retire at 65 and retain female work patterns (pesos)	7,330	10,016	17,717	29,215	54,947
Women retire at 60 and adopt men's working patterns (pesos)	8,504	11,611	18,701	25,315	43,017
Women at retire 60 and earn same earnings as men (pesos)	7,601	10,624	16,144	25,972	67,484
Men retire at 65	15,579	21,963	32,354	43,971	100,470
Eliminating the gender difference in fund accumulation (percent)					
Women's fund at 60/men's fund at age 65	36.02	35.66	42.68	51.86	41.72
(1) Effect of raising retirement age/men's fund at age 65	11.03	9.95	12.08	14.59	12.97
(2) Effect of increasing work experience/men's fund at 65	18.56	17.21	15.12	5.72	1.09
(3) Effect of eliminating income differentials/men's fund at 65	12.77	12.72	7.22	7.21	25.45
(4) Effect of the interaction of changes (1), (2), and (3)	21.62	24.47	22.90	20.63	18.77

Note: Estimated fund assumes 5 percent return and 2 percent secular growth in earnings.
Source: Author's estimates based on CASEN 94 data.

annuity derived exclusively from her accumulated funds would be necessarily smaller than that of a man derived from his accumulated fund.

To calculate the annuities from the estimated fund, I assume that the typical man is married to a woman three years younger than he is (this assumption is consistent with CASEN data). This man retires at age 65 with a life expectancy of 15 years. Because his wife is expected to survive him by six years he is required to provide for six years of survivor's pension, at 60 percent of his own pension. Chilean law requires retiring married men to put aside funds to cover pensions for their widows and surviving children (the amount required to comply with this regulation is determined through a private contract between the retiree and an insurance company). The law does not require retiring married women to provide for their surviving husbands, unless the husband is handicapped. I assume that there are no surviving minors, that men reserve part of their funds to provide for their widows, and that women convert their entire fund to an annuity. I also assume that insurance companies are allowed to use different survival tables for men and women. If the resulting annuity is smaller than the minimum pension guarantee, the estimated value is increased to the minimum pension.

Based on the calculated annuity, women's replacement ratios (the estimated annuity divided by the reference salary) are almost 60 percent of men's. These differences in replacement rates are smaller than the measured differences in the accumulated funds and the annuities, however, for three reasons. First, several categories of women qualify for the minimum pension, which raises the annuity above the level supported by own funds. Second, replacement ratios are calculated as the ratio of the monthly annuity over the reference salary, which is the average tax base of the last 10 calendar years of work divided by 12. The reference period corresponds to 120 calendar months. To the extent that the typical man or women works less than 120 calendar months during the reference period, the estimated reference salary is lower than the estimated average income for the same reference period. During the 10 years that precede the minimum pensionable age, women have an average work accumulation that is significantly lower than that of men. This causes the gender differential in reference earnings to be larger than the gender differential in earnings. Third, if the denominator in women's replacement ratio is relatively low, the resulting replacement ratio for women is relatively high.

The Issue of Retirement Age

A significant fraction of men and women who have reached retirement age choose not to claim benefits. Thus the retirement age operates as an option that people take when it is convenient for them. If women remain in the labor force beyond age 60, they can add to their fund accumulation, and they are more likely to qualify for the minimum pension on the grounds of years of contributions. The impact on annuities is positive because the accumulated fund is larger and the number of years to be covered by the annuity is smaller.

Under the low-returns scenario, women with less than complete secondary schooling have no gains from delaying claims because they qualify for the minimum pension guarantee. In contrast, women who complete secondary schooling stand to increase their annuity by 10 percent if they postpone claims, and women with higher levels of schooling stand to gain much more. Under the high-returns scenario, postponing claims increases the annuity about 50 percent for all groups. Whether an additional year of work past age 60 increases or decreases these women's welfare depends on individual preferences. One can say only that allowing people to draw a pension early or late is a superior alternative to imposing an age requirement before claiming benefits or forcing retirement at a given age.

The Impact of Reform

Under the old pension system rules, the monthly retirement benefit was equal to zero if the affiliate made less than 10 years' worth of contributions or the maximum of

$$0.50BS + 0.01BS(W - 500) / 50 \text{ or } 0.70BS$$

where BS (the base salary) equals the sum of total taxable earnings of the previous five years divided by 60, indexing the last three years, and W equals the total number of weeks of accumulated experience (beyond 520).[6] Men could retire at age 65 and women at 60. (In 1979, the retirement age for women was raised from age 55 to 60 in an effort to contain the system's growing deficit.) Benefits included survivor's pensions equivalent to 50 percent of the pension of the originator for widows and 20 percent of the mean salary per child. Men typically work 40 years, and a man with 40 years of contributions was very close to the maximum replacement rate of 70 percent. Therefore, men's pensions and widow's pensions were generally capped.

Women with more than 10 years of contributions got a very good deal under the old system rules. They could retire 5 years earlier than men, receive a benefit based on their last 5 years of earnings, and receive a 60 percent replacement rate based on a typical life time experience of just 20–30 years of work. Women with less than 10 years of contributions faced no incentive to participate in the system because they did not qualify for benefits.

The deal was not so good for married women. Under the rules of the old social security system, which covered the majority of the currently retired population, women had to choose between retirement income and pension. That is, if they were eligible for benefits from their own working years and were also eligible for a widow's pension, they could not receive both two sets of benefits and had to choose the better of the two.[7]

6. Because rules vary across funds, I used the social security system rules. The social security system represented more than 60 percent of contributors in 1980.

7. Art. 7, Law 10.383

Under the rules of the new social security system, benefits are a function of the accumulation of funds. There is no minimum number of years of contributions required to obtain a pension, as there was in the old system, which required at least 10 years. Contributions accrue to the accumulated fund independently of the timing of labor force participation and independently of the periodicity of income-generating activities. In fact, women who make contributions early in their careers get credit for the compound interest associated with those early contributions.

On the one hand, the new system pays benefits as a function of contributions, which tends to lower some women's benefits and raise the benefits of single men to the extent that they are not subject to a maximum benefit. On the other hand, the new system includes a minimum pension guaranteed, which raises the benefits for women with low levels of schooling significantly above the actuarially fair levels. Moreover, a married man is required to fund his wife's pension as a function of her probability of survival, lowering the benefits of married men relative to single men.

Table 4 provides estimates of social security–related incomes for elderly men and women in each of the schooling categories. To estimate widows' pensions, I assume that married couples belong to the same schooling category.[8] The calculations highlight the complex effects of the system's reform. In particular, they suggest systematic effects by marital status and level of schooling. The calculations at the bottom of table 4 assume full indexation. However, because the old system was not fully indexed, the benefit estimates for the old system represent an upper bound. To give an idea of the degree of overestimation of these benefits, the following calculation is of interest. If benefits are maintained at the nominal level shown in table 4 and there is a 10 percent annual inflation, in the absence of indexation real benefits fall to less than 50 percent of their initial value after 10 years, to less than 20 percent after 17 years, and to a little more than 10 percent after 23 years.

Direct comparisons of benefits before and after reform do not take into account the degree of sustainability of the systems, in particular the fact that the old system was unable to deliver the promised benefits under its formulas. They

8. These estimates can be compared with those provided by Baeza Valdés and Burger Torres (1995), whose estimates are based on actual retirement cases. Using a sample of 4,064 people who retired under the new system, they estimate that the average replacement rate is 78 percent. The highest (relative) pensions were obtained by people who opted for early retirement, with a replacement rate of 82 percent under programmed retirement. Baeza Valdés and Burger Torres attribute this result to the fact that only those who enjoyed rapid accumulation of funds—mostly by making voluntary contributions—can opt for early retirement. Through December 1997 average old-age pensions under the capitalization system were 39 percent higher than average pensions under the old pay-as-you-go regime. Disability pensions under the new system were 61 percent higher than under the previous regime. Overall, replacement rates have been high—indeed, higher than in most industrial countries (see Davis 1998 and Gruber and Wise 1999). Naturally, because the Chilean system is a defined-contribution system, there are no assurances that the replacement rates observed until now will be maintained in the future.

TABLE 4. Estimated Retirement Incomes under Old and New Systems (1994 pesos/month)

	Incomplete primary	Incomplete secondary	Secondary	Up to 4 years postsecondary	More than 4 years postsecondary
Married men, retiring at 65					
New system	97,917	138,043	203,356	276,368	631,482
Old system[a]	86,775	136,776	214,990	335,491	764,117
Women, new system					
Own pension adjusted for minimum pension age 60–77[b]	37,738	43,679	77,010	127,169	233,788
Widow's pension, age 78–83	64,331	90,693	133,604	181,571	414,880
Own and widow's pension, age 78–83[c]	64,331	133,995	209,950	307,642	646,650
Old system					
Own pension, working women age 60–77[d]	28,508	48,661	116,798	185,780	333,517
Widow's pension, age 78–83	43,388	68,388	107,495	167,746	382,059
Own or widow's pension, age 78–83	43,388	68,388	116,798	185,780	382,059

Note· Data are for urban areas only. Calculations are based on 5 percent return on funds and 2 percent secular income growth. The calculation of benefits under the old system's rule is based on the concept of a base salary, the average amount earned during the 10 years preceding pension benefits.

[a]These benefits are at the maximum (70 percent of the base salary).

[b]The estimated annuity for the typical woman in the lowest schooling categories falls below the minimum pension. The estimated income is replaced by the minimum pension ($37,738)

[c]Because the widow's pension is significantly above the minimum pension, it is assumed that the beneficiary would stop receiving the minimum pension (and her own funds would be exhausted).

[d]Estimated monthly income under the old social security system starts at age 60. Old system benefits for women in the two upper schooling groups are at the maximum. Benefits are very close to the maximum for the lower schooling categories.

Source· Author estimates based on CASEN 94 data.

also ignore the significant reduction in contribution rates. In order to keep the system solvent, contributions would have to be raised or benefits cut. If the relative benefits of men and women were not affected, gender ratios of benefits before and after the reform can be compared.

Table 5 reports the gender ratio of accumulated contributions and the gender ratio of the present value of benefits for various categories of men and women. The first row in the top panel shows that contributions by women who completed secondary schooling were 43 percent those of men with the same level of schooling. The ratio of accumulated contributions of women relative to men is always less than 1 and typically below 50 percent. (Decomposition of the factors contributing factors to this differential—earlier retirement, lower earnings, less attachment to the labor force—is provided in table 3.)

TABLE 5. Gender Differences in Contributions and Benefits by Schooling, Marital Status, and Pension System

Item	Incomplete primary	Incomplete secondary	Secondary	Up to 4 years postsecondary	More than 4 years postsecondary
Ratio of accumulated contributions					
Women/men	0.44	0.43	0.52	0.63	0.51
Married men/single men	1.00	1.00	1.00	1.00	1.00
Ratio of present value of benefits, old system					
Single women/single men	0.53	0.58	0.88	0.90	0.71
Married women/single men	0.58	0.62	0 88	0.90	0.73
Nonworking married women/single men	0.19	0.19	0.19	0.19	0.19
Married men/single men	1.00	1.00	1.00	1.00	1.00
Ratio of present value of benefits, new system					
Single women/single men	0.54	0.45	0.54	0.65	0.52
Married women/single men	0.68	0.60	0.69	0.81	0.68
Nonworking married women/single men	0.13	0.13	0.13	0.13	0.13
Married men/single men	0.87	0.87	0.87	0.87	0.87

Note: Data are for urban areas only Calculations are based on 5 percent return on funds and 2 percent secular income growth. Calculation assumes women retire at age 60 and men retire at age 65.

Source: Author's estimates based on CASEN 94 data.

Under the old system's rules, working women with secondary school education drew 88 percent of single men's benefits (table 5). Under the old system there was no strict link between contributions and benefits. The gender disparity between contribution ratios and benefits ratios indicates a significant pro-women bias in the allocation of benefits. The pro-women bias was also regressive: For women with secondary education, the ratio of benefits to contributions was 30 percentage points higher for women than for men, whereas for women with less than secondary schooling, the difference fell to less than 20 percentage points.

Under the new system, single and married working women with secondary school draw 54 percent or 69 percent, respectively, of the benefits that single men obtain (table 5). Benefits are calculated differently for single and married men because married men are required to provide survivor's benefits. Single women are beneficiaries if they have made contributions; married women are beneficiaries if they are either married to a contributor or made contributions themselves.

Contributions and benefits are closely linked, except for two factors. The joint annuities provision requires the redistribution of benefits from husbands to wives; the minimum pension guarantee provides additional benefits from sources outside the system to those who would otherwise draw pensions below a minimum. Single women's benefit ratios are just below the corresponding contributions ratios (see the bottom of table 5), except for the case of low schooling categories, where the minimum pension raises benefits significantly higher than contributions. (The difference between the ratio of contributions and the ratio of benefits in a defined contribution system comes from the cost of annuities. In Chile contributors pay a cost for transforming their accumulated fund into an indexed annuity. This cost is assumed here to be higher for women because they take their annuity for a longer period.) Married women have benefits ratios that exceed their contribution ratios because they receive widows' pensions. For the same reason, for a given ratio of contributions, the ratio of benefits of married men to single men is lower than 1.

Consistent with the fact that the old system was unsustainable, all benefit ratios presented in table 5 are higher than or equal to their corresponding contribution ratios. This includes noncontributing married women who receive a benefit that does not have a counterpart in reduced benefits for married men. This means that the old system contained a significant labor tax component, a significant inflation tax component, or both. In fact, inflation and the system's deficit were part of the picture at the time of reform, which introduces another problem in comparing the benefits between the old and the new systems.

Another feature of the benefit structure of the old system is that women married to men with high earnings received a relatively larger unfounded benefit from the system than women married to men who earned less. This adds a regressive dimension to the old system biases, because contributors finance widows' pensions, which are more generous for widows of men who had had high earnings.

The old system appears to have had a significant bias in favor of some women and their families, particularly married noncontributor women and married contributors with more than secondary schooling. However, inflation reduced the real value of benefits relative to contributions for all groups. Did the gender bias remain? The answer is empirical because it depends on the impact of inflation on real benefits, which is a function of indexation. Because the system was characterized by a notorious absence of indexation protections and women have longer life expectancies than men, one can say that the real gender bias was smaller than the nominal one. The new system reversed the regressive nature of the gender bias, as it forced each married man to make provisions for his own widow. In addition, it removed the bias against single men, whose funds no longer helped finance widows' pensions.

Who Benefits from Guaranteed Pensions?

In principle, the state guarantees minimum old age, invalidity, and survival pension benefits to affiliates and their beneficiaries, as long as they are poor and have made contributions for at least 20 years. According to the law, "No one can obtain the state subsidy if the sum of all individual incomes from pensions, rents, and taxable earnings is equal to or higher than the minimum pension."[9] Therefore, unlike access to earned benefits through contributions, access to the guaranteed minimum pensions can be taken away if income-generating conditions change. In fact, qualifying affiliates who also receive the PASIS benefit must give up that pension as soon as the guaranteed minimum benefit is activated.[10]

In practice, the means-testing procedures have not been fully incorporated, and qualification toward the minimum pension is mainly a function of accumulated funds and life expectancy. Putting aside the means testing, an accurate estimate of the number of affiliates who would qualify for the minimum pension at retirement, or for invalidity or survivor's benefits while active, requires longitudinal data on individual contributions.

The typical woman in the lowest schooling category would qualify for the minimum pension under low and high returns; under low returns the typical women in the next two schooling categories would also qualify for the state guarantee. Typical men in the lowest schooling category would qualify for the minimum pension only under low returns, although their accumulated funds are just under the minimum necessary to generate the minimum pension. The mini-

9. Art. 80, DL 3,500

10. Accidents on the job are covered by insurance, which pays out in proportion to reference salaries. The state guarantee affects people who earn very low salaries, who work few hours or contribute sporadically, or who become incapacitated or die early in their career, leaving a large number of legal survivors. The minimum invalidity pension is paid to affiliates who are declared legally incapacitated, do not qualify for the minimum pension, and fulfill minimum qualifying conditions. There is also a minimum survivor's pension (60 percent of the minimum pension) paid to legal survivors of affiliates who fulfill minimum qualifying conditions.

mum pension program is thus expected to favor women, as Wagner (1991) and Zurita (1994) have noted.

This exercise in comparative static must be viewed with caution, however, because the relation between earnings and accumulation is affected by the individual's contribution density and the system's rate of return. While the rate of return is a parameter of the system, the density of contributions is a behavioral variable. The current system's rules discourage social security contributions after 20 years worth of accumulated contributions by people whose taxable earnings are low enough to qualify them for the minimum pension. Thus one should expect to observe many people with earnings around the minimum income applying for pensions with exactly 20 years of contributions.[11]

One of the weaknesses of the current rules is that the minimum contribution (on minimum incomes) is not enforced. In particular, if a worker reports part-time income, the minimum contribution becomes in effect lower than the minimum wage. If authorities counted part-time employment (relative to the minimum legal contribution) as partial time, the possibility of making contributions below the legal minimum would be eliminated. The implication is that the number of calendar years of contributions needed to qualify for the minimum pension will be more than 20 for people who work part-time. The effective years of contributions, measured in full-time equivalent minimum earnings, to qualify for the minimum pension will still be 20 years. In addition, authorities can restrict access to the minimum pension by making the program a truly means-tested program. At the very least, information regarding access to widows' pensions should be taken into account to examine eligibility toward continued minimum pension benefits for women.

V. Summary and Conclusions

The Chilean pension reform benefited contributors on three fronts: It reduced contributions, it established indexation of benefits, and it made the system sustainable by tying benefits directly to contributions. The reform established a distributive pillar funded directly by the government budget. All these elements reinforce the effect of reducing the tax on labor, encouraging labor force participation and employment. At the same time, the direct link between contributions and benefits required the elimination of cross-subsidies within the system, a source of complex effects on the relative position of women.

It is argued here that the tax reduction effect of social security reform was more pronounced on women. First, under the new system there is no minimum

11 The estimates assume that all workers in a given schooling group are identical in terms of their work patterns. In fact, this is not the case; there is a distribution around the mean. Because some workers have less than 20 years of work, they would not qualify for the minimum pension. Some workers with more than 20 years of work and earnings above the mean may not qualify if their funds are sufficient to fund an annuity above the minimum pension.

level of contributions to obtain a pension (under the old system contributors with less than 10 years of contributions did not receive pension benefits). Second, the new system allows widows to keep their own pension benefits in addition to their widow's pension, restoring the marginal benefit of own contributions for working women. Third, the reform gives more weight to early years of contributions (as a result of compound interest), rather than the heavy weight given to the last five years in determining the pensionable income under the old system. This change favors women relative to men because women are more likely to hold paid jobs when young and to drop out of the labor force later. Moreover, even if women maintain a significant attachment to the labor force, they tend to have flatter age-earnings profiles than men.

The new system removed three biases associated with funding of survivor's pensions. In Chile's defined-contribution system, survivors' pensions are funded directly by contributors. In the traditional pay-as-you-go system, survivors' pensions were funded from the general system funds. Thus in the traditional system when a rich old man married a young woman a few years before retiring, he would draw a generous pension until his death and bequeath to his widow a generous pension funded from all contributors in the system. In contrast, in the new system, the old man would draw a smaller pension so that after his death the remaining amount would fund a proportional pension for his young widow, who is expected to live many years. Therefore, the new system removed the bias that favored married men (who did not have to make provisions for their widows' pensions) and the bias in favor of widows of rich men who obtained relatively generous pensions financed by all contributors. In addition, according to the rules that apply to the majority of beneficiaries of the old system, if a widow receives her own benefits, she has to choose between those and her widow's pension. In contrast, own pension and survivor's pensions are complementary in the new system. Thus the reform eliminated a bias against married working women.

But the new system also contains its own equity–efficiency tradeoffs, particularly with respect to women. The minimum pension mainly benefits women, given their low rates of pay and limited years of contributions. This study focuses on the work patterns of women affiliates, a subsample of women with a stronger attachment to the labor force. In this group, the typical woman now works about 20 years. Clearly, the new rules will encourage everyone in this group to work to accumulate 20 years of paid work to qualify for the guarantee. The guarantee targets low earners rather than middle-class women. However, once low-earning women (that is, women without substantial postsecondary education) qualify for the minimum pension, they get little if any additional benefit for incremental years of contributions. In effect, they are subject to a heavy implicit marginal tax rate on their labor. The minimum pension effectively becomes a ceiling as well as a floor. Thus the new policy is well designed to keep working women out of poverty in their old age, given their current labor market behavior, but it also maintains that behavior, with transitory labor market attachment, for women with limited education. Although this may be an improvement over the previ-

ous policy, policymakers may wish to reevaluate this guarantee and tie it more continuously to years of contributions to provide a safety net with even more positive incentive effects.

Perhaps one of the oversights of the Chilean reform was to set women's pension age at 60. Given the differences in longevity, women would have to save more than men do to obtain the same retirement incomes. In fact, contributing women tend to accumulate less than contributing men because of lower attachment to the labor force and lower earnings than men. Therefore, if not for the minimum pension, women's pensions have to be lower than men's. The expectation of a survivor's pension makes the combined pension benefits of married women higher than that derived solely from own funds. But not all women who work for pay are married or will inherit a pension to complement their incomes in old age.

Overall, the new system is more fiscally sustainable and creates an incentive for greater labor force participation. Women who never enter the labor force are in a more vulnerable position in old age, and this vulnerability is likely to increase as they age. The fact that the Chilean social security reform improved women's incentives to work for pay offers hope for behavioral changes that would reduce women's risk of poverty in old age.

One important lesson to be drawn from this research is that the different work histories of men and women should affect the design of the public pillar and its eligibility requirements. The number of years chosen as a qualifying condition for the minimum pension guarantee is a critical determinant of the gender effects of reform. In Mexico, where affiliates need to contribute for 25 years to qualify for the pubic benefit, and Argentina, where 30 years of contributions are required, the gender impact is probably different, with men instead of women benefiting disproportionately.

References

Baeza Valdés, Sergio, and Raúl Burger Torres. 1995. "Calidad de la Pensiones del Sistema Privado Chileno." In S. Baeza and F. Margozzini, eds., *Quince Años Después: Una Mirada al Sistema Privado de Pensiones*. Santiago: Centro de Estudios Públicos.

Davis, E. Phillip. 1998. "Pensions in the Corporate Sector," in Siebert, ed., *Redesigning Social Security*. Tubingen: Mohr Siebeck.

Edwards, Alejandra Cox. 2000. "A Close Look at Living Standards of Chilean Elderly Men and Women." California State University–Long Beach, Department of Economics.

———. 2001. "Social Security Reform and Women's Pensions." Working Article Series 17, for World Bank Policy Research Report Gender and Development. World Bank, Washington, D.C.

Edwards, Sebastian, and Alejandra Cox Edwards. 2002. "Social Security Privatization Reform and Labor Markets: The Case of Chile." *Economic Development and Cultural Change* 50(3):465–89.

Gruber, Jonathan, and David A. Wise, eds. 1999. *Social Security and Retirement around the World*. Chicago: University of Chicago Press for National Bureau of Economic Research.

SAFP (Superintendencia de Administradora de Fondos de Pensiones). Various years. *Boletín Estadístico* (various issues). Santiago.

Wagner, Gert. 1983. "Estudio de la Reforma Previsional: Previsión y Reforma, efectos en la industria y en el país." Universidad Católica, Instituto de Economía (Mayo).

———. 1991. "La Seguridad Social y el Programa de Pensión Mínima Garantizada." *Estudios de Economía* 18(1):35–91.

Zurita, Salvador. 1994. "Minimum Pension Insurance in the Chilean Pension System." *Revista de Análisis Económico* 9(1):105–26.

THE WORLD BANK ECONOMIC REVIEW, VOL. 16, NO. 3 345–373

Low Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development

Stephan Klasen

Using cross-country and panel regressions, this article investigates how gender inequality in education affects long-term economic growth. Such inequality is found to have an effect on economic growth that is robust to changes in specifications and controls for potential endogeneities. The results suggest that gender inequality in education directly affects economic growth by lowering the average level of human capital. In addition, growth is indirectly affected through the impact of gender inequality on investment and population growth. Some 0.4–0.9 percentage points of differences in annual per capita growth rates between East Asia and Sub-Saharan Africa, South Asia, and the Middle East can be accounted for by differences in gender gaps in education between these regions.

Many developing countries exhibit considerable gender inequality in health, employment, and education. For example, girls and women in South Asia and China suffer from much higher mortality rates than do men—creating what Amartya Sen calls "missing women" (Klasen and Wink 2002, Sen 1989). Employment opportunities and pay also differ greatly by gender in most developing regions (as well as most industrial regions; see United Nations Development Programme [UNDP] 1995 and World Bank 2001). Finally, there are large gender discrepancies in education, particularly in South Asia, the Middle East and North Africa, and Sub-Saharan Africa.

When assessing the importance of these gender inequalities, one has to distinguish between intrinsic and instrumental concerns. If the concern is aggregate well-being—as measured by, for example, Sen's notion of "capabilities" (Sen

Stephan Klasen is with the Department of Economics, University of Munich. His e-mail address is klasen@lrz.uni-muenchen.de. The author is grateful to Jere Behrman, Chitra Bhanu, Mark Blackden, François Bourguignon, Lionel Demery, David Dollar, Bill Easterly, Diane Elson, Roberta Gatti, Beth King, Stephen Knowles, Andy Mason, Claudio Montenegro, Dorian Owen, Susan Razzazz, Lyn Squire, Martin Weale, Jeffrey Williamson, and participants at various seminars and workshops for helpful comments on earlier drafts of this article. The author also thanks Frank Hauser for excellent research assistance. World Bank funding from a research grant in support of this work is gratefully acknowledged

DOI: 10.1093/wber/lhf004

1999)—then longevity and education should be seen as crucial constituent elements. Given inequality aversion (or, equivalently, declining marginal social valuation of these achievements), gender inequality in these achievements will reduce aggregate well-being.

In addition, one may be concerned about gender equity as a development goal in its own right (apart from its benefits for other development goals)—as recognized by the Convention on the Elimination of All Forms of Discrimination against Women, which has been signed and ratified by 165 countries (UNDP 2000). If this is the concern, there is no need to do more than demonstrate inequity in a particular country, which would justify corrective action.

Apart from intrinsic problems of gender inequality, one may be concerned about instrumental effects of gender bias. Gender inequality may undermine a number of development goals. First, gender inequality in education and access to resources may prevent reductions in fertility and child mortality and expansions in education of the next generation. A large literature documents these linkages (Klasen 1999, Murthi, Guio, and Drèze 1995, Summers 1994, Thomas 1990, 1997, World Bank 2001). Thus gender bias in education may generate instrumental problems for development policymakers as it compromises progress on these other important development goals.

Second, gender inequality may reduce economic growth. This issue is important to the extent that economic growth advances well-being (as measured by such indicators as longevity, literacy, and poverty), though not all types of growth do so to the same extent (Dollar and Kraay 2000, Drèze and Sen 1989, Pritchett and Summers 1996, Ravallion 2001, UNDP 1996, World Bank 2000). Thus policies that advance economic growth (and do not impede other development goals) should be of great interest to policymakers.

This article focuses on the instrumental effect of gender inequality in education on economic growth.[1] Using cross-country regressions, it shows how gender bias in education reduces economic growth. This effect accounts for a considerable portion of the differences in growth between developing regions. In particular, South Asia, the Middle East, and Africa are held back by high gender inequality in education.

I. Previous Findings on Inequality in Education and Economic Growth

Recent years have seen renewed interest in the theoretical and empirical determinants of economic growth. On the theoretical front, Roemer (1986), Lucas (1988), and Barro and Sala-i-Martin (1995) emphasize the possibility of endogenous economic growth where growth is not constrained by diminishing returns

1. A longer, more detailed working paper (Klasen 1999) also considers the impact on growth of gender inequality in employment and examines the impact of gender inequality in education on fertility and child mortality.

to capital. These models stand in contrast to Solow (1956), which, by using a neoclassical production function (with diminishing returns to each input) and exogenous savings and population growth, suggests convergence of per capita incomes—conditional on exogenous savings and population growth rates. Many growth models also emphasize the importance of human capital accumulation. Human capital can be included in the traditional Solow model and still yield conditional convergence. It can also be incorporated into (and indeed, be a major reason behind) endogenous growth models (Barro and Sala-i-Martin 1995, Mankiw, Roemer, and Weil 1992).

Few growth models explicitly consider the impact of gender inequality in education. Knowles, Lorgelly, and Owen (2002) extend the Solow model by considering male and female human capital as separate, and thus imperfectly substitutable, factors of production. Given diminishing returns to each factor, a more balanced distribution of education among males and females would lead to higher steady-state per capita income.

Lagerlof (1999) and Galor and Weil (1996) examine the links between gender inequality in education or earnings on fertility and economic growth in an overlapping generations framework. Lagerlöf shows that initial gender inequality in education can result in high fertility, low economic growth, and continued gender inequality in education, thus creating a poverty trap that justifies public intervention.[2] Similarly, Galor and Weil postulate that economic growth narrows the gender gap in earnings, lowering fertility and advancing economic growth. Conversely, low-income countries, who often have high gender gaps in earnings, will suffer from high fertility and low economic growth, which will perpetuate the gender earnings gap and generate a similar poverty trap. In both models gender gaps in education and earnings reduce economic growth mainly through their demographic effects.

On the empirical front, the development of international panel datasets has, for the first time, allowed sophisticated time-series, cross-section, and panel analysis of the determinants of long-run growth (e.g., Asian Development Bank [ADB] 1997, Barro 1991, Barro and Sala-i-Martin 1995, Mankiw, Roemer, and Weil 1992). These empirical studies are usually based on estimating the transitional dynamics to a steady state in a Solow framework and so examine the issue of convergence as well as the importance of savings, population growth, and human capital accumulation. They often add largely ad hoc specifications to proxy for country differences in the technological shift parameter (broadly conceived) of the production function. These shift parameters affect steady-state gross

2. Lagerlof's model assumes that parents aim to maximize the household incomes of their children. If there are gender gaps in education, it may be optimal for parents to focus education investments on sons because daughters are likely to marry an educated man, whereas sons are likely to marry an uneducated woman. These self-perpetuating gender gaps in education lead to high fertility because the opportunity cost of time is low for uneducated women. High fertility results in low investments in each child, which can mire an economy in a poverty trap. Public action to reduce the gender gap can break this self-perpetuating cycle and help move the economy out of a poverty trap.

domestic product (GDP) levels in conditional convergence models and long-term growth rates in endogenous growth models. Factors examined in this context include the level of corruption, political instability, ethnolinguistic fractionalization, openness to trade, geographic and climatic constraints, the quality of public services and institutions, and the dependency burden (ADB 1997, Barro 1991, Bloom and Williamson 1998, Collier and Gunning 1999, Easterly and Levine 1997, Sachs and Warner 1995).

These empirical analyses differ not only in their inclusion of parameters thought to affect steady-state GDP but also in their approach to treating proximate determinants of economic growth—especially the investment rate. Some studies include the investment rate as an explanatory variable (Barro and Sala-i-Martin 1995, Mankiw, Roemer, and Weil 1992). Others leave it out in the belief that it is determined by other regressors in the equations, such as the level of human capital or the quality of institutions (ADB 1997, Barro 1991, Bloom and Williamson 1998). Thus the second type of analysis has estimated "reduced form" equations in which proximate determinants of economic growth (especially the investment rate) are omitted from regressions to measure the total effect of independent variables.

Taylor (1998) presents results on both types of regressions. He first regresses growth rates on a number of structural factors derived from the Solow growth model (investment rate, population growth rate, initial income, and some shift parameters). Then he regresses each of those factors on a range of other determinants, such as policy distortions and political, social, and demographic variables.[3] Thus he aims to show the direct and indirect effects on growth of some policy variables and to better understand the process that leads to poor economic growth. For example, the population growth rate may affect economic growth directly, but it may also affect it through its impact on the investment rate or human capital accumulation.[4]

Although their results differ, the studies discussed generally show that there is evidence of conditional convergence, that initial and subsequent investments in human capital are associated with higher growth, that population growth dampens economic growth, that openness increases economic growth, and that high government consumption, political instability, and ethnic diversity reduce economic growth (Barro 1991, Bloom and Williamson 1998, Easterly and Levine 1997). Being landlocked or located in the tropics, having a small coastline, and having a large natural resource endowment also appear to dampen growth (ADB 1997, Sachs and Warner 1997).

3. This procedure can be done using ordinary least squares and yields consistent estimates because the system of equations being estimated is recursive, not simultaneous. Thus one can first consistently estimate the structural equations determining the endogenous variables, then estimate the growth regression using those endogenous variables. But care must be taken in interpreting the results, because direct effects must be distinguished from indirect effects.

4. In addition, some variables may influence economic growth only through their impact on structural parameters Taylor (1998) finds, for example, that government spending affects GDP growth not directly, but only through its impact on investment.

Few studies have explicitly considered the impact on growth of gender inequality in education. Barro and Lee (1994) and Barro and Sala-i-Martin (1995) report the "puzzling" finding that in growth regressions that include male and female years of schooling the coefficient for female primary and secondary schooling is negative. They suggest that a large gap in male and female schooling may indicate backwardness and so may be associated with lower economic growth.

But there are reasons to question that "puzzling" finding. Dollar and Gatti (1999) show that it disappears once a dummy variable for Latin America is included, suggesting that the finding may be due to the combination of low growth and comparatively high female education in Latin America.[5]

The finding may also be related to multicollinearity. In most countries male and female education levels are closely correlated, making it difficult to identify empirically their individual effects. The correlation coefficient between male and female years of schooling and similar attainment measures (such as the share of adults with secondary education) is consistently above 0.9 for the large sample of countries considered by Barro and colleagues and in this study. Large standard errors for male and female education and the sudden reversal of this finding in different specifications is further evidence of this problem (Forbes, 2000, Knowles, Lorgelly, and Owen 2002, Lorgelly and Owen 1999). To avoid these pitfalls, this article adjusts the specification of education variables (see following discussion) and includes regional dummy variables.

Hill and King (1995) also study the effect on income of gender inequality in education. Instead of trying to account for GDP growth, they relate gender inequality in education to GDP levels. They find that a low female-male enrollment ratio is associated with lower GDP per capita, over and above the impact of female education levels on GDP per capita.

Knowles and others (2002) estimate the impact of gender inequality in education on GDP per capita using an explicit Solow framework, treating male and female education as separate factors of production. They find that gender inequality in education significantly reduces GDP per capita.[6]

This article differs from Hill and King and from Knowles, Lorgelly, and Owen in that it tries to explain differences in long-term growth rates of GDP rather than in levels of GDP per capita. It does so by including other standard regressors from the empirical growth literature. In addition, it differs from Hill and King in that it uses a broader, longer dataset and a more reliable measure of human capital.

5 Moreover, as Knowles, Lorgelly, and Owen (2002) argue, the use of base-period values for human capital in Barro and colleagues' growth regressions contributes to this effect because high-growth East Asian economies had large initial gender gaps in 1960, so the coefficient on female education might pick up East Asia's initial gaps as growth-enhancing. Inclusion of regional dummy variables should deal with this problem as well. In addition, Lorgelly and Owen (1999) suggest that this problem may be partly due to the impact of outliers.

6. Knowles, Lorgelly, and Owen (2002) refer to Klasen (1999) and specifically include in their specifications this article's parametrization of measuring the impact on growth of gender inequality. They find that this article's parametrization best deals with the problem of collinearity and confirm its main findings.

Dollar and Gatti (1999) also examine the relationship between growth and gender inequality in education. They try to explain five-year growth intervals (between 1975 and 1990) and to control for the possible endogeneity between growth and education using instrumental variable estimation. In contrast to Barro and colleagues, they find that female secondary education (as measured by the share of female adults with some secondary education) is positively associated with growth, whereas male secondary education is negatively associated with growth. In the full sample both effects are insignificant. But in countries with low female education, increasing it has little effect on economic growth—and in countries with high female education, increasing it significantly boosts economic growth.[7]

This study differs from Dollar and Gatti in that it uses a longer growth interval (based on the presumption that human capital pays off only over the long term), a longer time period (beginning in 1960 rather than 1975), and a different measure of human capital. In addition, it addresses the multicollinearity problem. The following discussion reconciles the different findings between Dollar and Gatti and this analysis.

Finally, an extensive literature shows that gender inequality in education contributes to higher fertility and child mortality (e.g., Hill and King 1995, Klasen 1999, Murthi, Guio, and Drèze 1995, Schultz 1994, World Bank 2001). Combined with the negative effects that high fertility and population growth have on economic growth,[8] the effects that gender inequality in education have on fertility and child mortality may point to an indirect link between gender bias in education and economic growth. Thus regressions that include the fertility rate or population growth as a regressor may understate the total effect of gender bias in education, making it important to explicitly model these indirect links.[9]

II. Theoretical Links between Gender-based Education Inequality and Economic Growth

Based on the previous discussion and related theoretical considerations, the following causal links between gender inequality in education and economic growth might exist.

7. Forbes (2000) and Caselli, Esquivel, and Lefort (1996) use GMM estimators, a version of panel data analysis, and also find a positive effect of female education and a negative effect of male education on economic growth. Both analyses are focused on other questions and only use a very small number of covariates and, in the case of Forbes, a nonstandard income variable (non-PPP adjusted income). Their estimates might therefore partly be due to these shortcomings.

8. For example, a study of the Asian Development Bank (ADB) finds that high population growth depressed annual per capita growth in Sub-Saharan Africa by 0.7 percentage point between 1965 and 1990. This factor alone accounted for about 15 percent of the difference in growth performance between Sub-Saharan Africa and Southeast Asia (ADB 1997). Similarly, Bloom and Williamson (1998) find that East Asia's early fertility transition was an important factor in the region's economic success.

9. The same might be true, to a lesser extent, for growth regressions that include child mortality or life expectancy.

Lower Average Human Capital

Assuming that boys and girls have a similar distribution of innate abilities and that children with more abilities are more likely to receive education, gender inequality in education means that less able boys than girls have the chance to be educated. As a result, the average innate ability of educated children is lower than it would be if boys and girls had equal education opportunities. Assuming that the amount of human capital of a person is the outcome of a combination of innate abilities and education, gender inequality in education would therefore lower the average level of human capital in the economy and therefore slow economic growth. For the same reason, such gender inequality would lower the impact of male education on economic growth and raises the impact of female education (Dollar and Gatti 1999, Knowles, Lorgelly, and Owen 2002).

Empirically, this lower human capital should directly reduce economic growth. This effect alone could shrink annual per capita growth by 0.3 percentage point in countries where gender inequality in education is similar to current levels in Africa, compared to a situation with no gender inequality in education.[10] It could also reduce the investment rate and thus indirectly reduce growth, because countries with lower human capital have smaller returns on investments.

A similar lowering of human capital occurs if male and female human capital are considered imperfect substitutes and there are declining marginal returns to education (Knowles, Lorgelly, and Owen 2002). In such cases diminishing returns to higher male education (rather than selection of less able males) lower the average level of human capital and thus economic growth.

All these effects are supported empirically. In many developing economies marginal returns to education are higher for girls than for boys, probably because of the selection effect and declining marginal returns to education (Alderman and others 1995, 1996, Hill and King 1995, World Bank 2001). Similarly, there is considerable evidence for the imperfect substitutability of male and female labor in many settings, and simulation studies have shown that a more equal allocation of male and female labor among industries would boost economic growth (Tzannatos 1999, World Bank 2001).

Wage Discrimination and Female Employment

In most countries women experience wage discrimination in formal employment. This discrimination shows up in earnings regressions as the unexplained portion of the female–male wage gap (Horton 1999, Tzannatos 1999, World Bank 2001). If women have enough education to participate effectively in the

10. The calculations assume that innate abilities are normally distributed and compares two possible distributions of the student population In one, half of an age cohort gets educated, and half are male and half are female. In the other, half of a cohort gets educated, and 70 percent are male and 30 percent are female. Average human capital would be 12 percentage points lower in the second scenario. Using the regression coefficient on human capital from Mankiw, Roemer, and Weil (1992) would yield a 0 3-percentage-point difference in annual growth.

formal labor market, wage discrimination can boost investment in industries that employ female workers. Reducing gender inequality in education may enable employers to employ cheaper female workers—boosting investment and thus economic growth.

Again, there is some empirical support for this effect. In many countries a considerable portion of high growth has been based on the use of female workers in export-oriented manufacturing industries. In some export-oriented Asian economies, for example, female education improved rapidly while there were large wage gaps between women and men, favoring female employment and the development of female-intensive industries (Seguino 2000b, Standing 1999).[11]

Direct Externality Effects

Lower gender inequality in education means higher female education at each level of male education. Because female education is believed to have positive external effects on the quantity and quality of education for educated women's children (through the support and general environment that educated mothers can provide), lower gender inequality would therefore improve the human capital of the next generation, which should also promote economic growth (World Bank 2001).

Moreover, to the extent that similar education levels at the household level have positive external effects on the quality of education, reduced gender inequality may be one way to promote such effects. For example, equally educated siblings can strengthen one another's educational success through direct support and play inspired by schooling. Similarly, couples with similar education levels can promote one another's life-long learning.[12]

The higher human capital associated with such processes can increase economic growth directly by boosting worker productivity. But it can also have an indirect effect by increasing returns to physical investment, which raises investment rates and—through the effect of investment on economic growth—increases economic growth.

Indirect Externality Effects Operating through Demographic Effects

As already discussed, there is overwhelming evidence that higher female education, which would obtain as a result of lower gender inequality in education,

11 Seguino (2000b) finds that gender gaps in education reduced economic growth in a sample of export-oriented middle-income economies, whereas gender gaps in pay increased it. Over time this effect will probably erode because the increased demand for female employment will reduce gender gaps in pay, as has been shown in most developing economies—including many fast Asian economies (World Bank 2001, Tzannatos 1999, Horton 1999). But in practice the unexplained portion of the gender gaps has far from disappeared in these countries, so this erosion appears to take a long time. Moreover, in some mature export-oriented economies where the share of manufacturing is declining—for example, Hong Kong (China) and Taiwan (China)—the gender gap appears to have widened recently. For discussions, see Berik (2000) and Seguino (2000a).

12. Even if people prefer to marry people with similar education levels (as appears to be the case), gender inequality in education often forces educated men to marry uneducated women, preventing this spillover. For a discussion of these and related issues, see Baliga, Goyal, and Klasen (1999).

reduces fertility rates. Lower fertility could affect economic growth in four different ways. First, lower fertility reduces population growth and thus facilitates investment's being used for capital deepening (more capital per worker) rather than capital widening (equipping new workers with capital), which would promote economic growth.[13]

Second, reduced fertility lowers the dependency burden, increasing savings rates in an economy, which would increase growth.

Third, lower fertility will, for a limited period of time, increase the share of workers in the population. When a large number of workers enter the labor force as a result of previously high population growth, it increases the demand for investment in capital equipment and social overhead (such as housing). If this higher demand is met by increased domestic savings (as a result of the reduced dependency burden), increased capital inflows, or both, investment will expand—which should boost growth (Bloom and Williamson 1998). This effect operates mainly through the impact of population growth on investment and its impact on economic growth rather than by affecting growth directly.

Fourth, if growth in the labor force is absorbed through increased employment, per capita economic growth will rise even if wages and productivity remain the same. This is because more workers will be sharing their wages with fewer dependents, boosting average per capita income. These last two effects—referred to as a "demographic gift" by Bloom and Williamson (1998)—are temporary because after a few decades growth in the working-age population falls and the number of elderly rises, leading to a higher dependency burden. Still, this temporary effect is thought to have made a considerable contribution to rapid growth in East and Southeast Asia (ADB 1997, Bloom and Williamson 1998, Young 1995; see also following discussion).[14]

III. Data, Measurements, and Specifications

A serious measurement problem arises when measuring the impact on economic growth of gender inequality in education. Because many subsistence, household, and reproductive activities are not captured by systems of national accounts, some estimates suggest that two-thirds of female economic activities go unrecorded in developing economies (compared with just one-quarter of male activities; see UNDP 1995, p. 89). Similarly, increases in the quantity and productivity of these

13. In a Solow framework, it would only boost economic growth in the transitional dynamics to a steady state (which can take a long time); in an endogenous growth framework, it could have permanent growth effects.

14. Lagerlof (1999) describes a fifth demographic effect that suggests an interaction between gender inequality in education, high fertility, and low investment in human capital, and so economic growth. In this case the impact of fertility mainly operates through human capital investments for the next generation. A related demographic effect may operate through the health of the next generation. Educated women have healthier children and so indirectly contribute to economic growth by producing healthier, more productive workers (Klasen 1999, Summers 1994).

activities are often recorded insufficiently or not at all (Waring 1988). Moreover, economic change may lead to lower measured levels of female economic activity (say, through increased informalization of female work or larger burdens on females as a result of smaller social safety nets), depressing measured economic growth (Palmer 1991).

Thus any study of the link between gender-based education inequality and economic growth will suffer from these data weaknesses. Any finding of the impact of gender inequality on economic growth may understate the true relationship—particularly if better female education increases not just female activities included in systems of national accounts but also activities not included.[15]

Bearing in mind that caveat, the following country data were used in this article's analysis:

- Data on incomes and growth are based on per capita incomes between 1960 and 1992 adjusted for purchasing power parity (PPP, expressed in constant 1985 U.S. dollars using the chain index), as reported in the Penn World Tables Mark 5.6 (Heston and Summers 1991, National Bureau for Economic Research [NBER] 1998).[16] An average compound growth rate, calculated for 1960–92, is the dependent variable. Data on investment rates, population growth, and openness (defined as exports and imports as a share of GDP) are also drawn from the Penn World Tables.
- Data on schooling are based on Barro and Lee (1996) and refer to the total years of schooling of the adult population.[17]

15. It is not safe to assume, however, that this is always the case. For example, it is possible to imagine scenarios in which reduced gender inequality increases women's activities included in systems of national accounts and reduces activities not included, resulting in a smaller than observed impact on total economic output. This may happen if, for example, women enter the labor force and hire child care for their children. Both moves increase measured output, but the hired child care is simply replacing previously unrecorded activities.

16. Data are not available for all countries in 1960 or through 1992. Thus the time period considered is shorter for some countries. This fact is taken into account in the calculation of average growth rates for all the variables.

17. In the regressions using data on initial schooling in 1960, adults are defined as people age 15 and older. In the panel regressions adults are defined as people age 25 and older. This difference does not have much effect on the results. The education measure differs from that in Dollar and Gatti (1999), who use the share of the adult population with some secondary education. There are advantages and disadvantages to both types of variables. Using years of schooling captures the average education of the adult population, whereas the other measure does not, for example, differentiate between adults with no education and adults with complete primary education (both are counted as having no secondary education). However, using years of schooling (implicitly) suggests that all adults achieved these average figures and that differences in years of schooling between people do not matter. The other measure, to a limited degree, takes into account such differences. The measure used by Dollar and Gatti—the share of adults who have exactly achieved some secondary education—is a bit peculiar because it excludes people who have achieved more than secondary education. In countries where a growing share of adults have more than secondary education (such as Canada or the United States), this measure will show stagnation and not reflect these further improvements. It would be preferable to use a measure that includes people with some secondary education as well as people with more education. See also the discussion that follows.

- Data on education spending are based on Barro and Lee (1994).
- Data on the working-age population are from version 3.0 of the database on Women's Indicators and Statistics compiled by the United Nations Children's Fund (UNICEF 1996).[18]
- Data on fertility, child mortality, and life expectancy are from World Bank (1993) and UNICEF (1992).

Because the focus here is the long-term economic growth rate, the first set of regressions treats 1960–92 as one observation and runs a cross-section regression, which is also the specification preferred by Knowles and others (2002). Like Taylor (1998), the analysis considers both the direct and indirect effects on growth of gender inequality in education. Thus a set of equations is estimated to capture both types of effects. In addition, following Bloom and Williamson (1998) and the earlier discussion, population growth and labor force growth are separated out. Population growth is expected to have a negative effect on growth, and labor force growth a positive effect. In the basic specification, the following equations are estimated.

$$(1) \quad g = \alpha_1 + \beta_1 Inv + \beta_2 Popgro + \beta_3 LFG + \beta_4 ED60 + \beta_5 GED + \beta_6 RED60 + \beta_7 RGED + \beta_8 X + \varepsilon.$$

$$(2) \quad Inv = \alpha_2 + \beta_9 Popgro + \beta_{10} LFG + \beta_{11} ED60 + \beta_{12} GED + \beta_{13} RED60 + \beta_{14} RGED + \beta_{15} X + \phi.$$

$$(3) \quad Popgro = \alpha_3 + \beta_{16} ED60 + \beta_{17} GED + \beta_{18} RED60 + \beta_{19} RGED + \beta_{20} X + \varphi.$$

$$(4) \quad LFG = \alpha_4 + \beta_{21} ED60 + \beta_{22} GED + \beta_{23} RED60 + \beta_{24} RGED + \beta_{25} X + \gamma.$$

$$(5) \quad g = \alpha_5 + \beta_{26} ED60 + \beta_{27} GED + \beta_{28} RED60 + \beta_{29} RGED + \beta_{30} X + \nu.$$

In these equations g is the average annual compounded rate of per capita income growth in 1960–92, *Inv* is the average annual rate of investment (as a percentage of GDP) in 1960–92, *Popgro* is the average annual compounded rate of population growth in 1960–92, *LFG* is the average annual compounded rate of growth in the labor force (ages 15–64) population in 1960–92, *ED60* is total years of schooling in 1960, *RED60* is the female–male ratio of total years of schooling of the adult population in 1960, *GED* is average annual absolute growth in total years of schooling between 1960 and 1990, *RGED* is the female–male ratio of the average annual absolute growth of total years of schooling between 1960 and 1990, and X's are other regressors typically included in cross-country growth regressions, including average openness, the log of income per capita in 1960 to test for conditional convergence, and regional dummy variables.[19]

18 Data for the working-age population cover only 1970–94. They were also compared with data from the World Bank, and the regression results did not depend on the data source.

19. In addition, a variety of policy variables were used, such as government consumption and institutional quality Although some were significant, they added little to the explanatory power of the regressions and were omitted in the final specification used here.

Equation (1) measures the direct effect on economic growth of education and gender bias in education. Because gender bias in education may affect investment, population growth, and (with a delay) labor force growth, equations (2), (3), and (4) measure the effects that education and gender bias in education have on these variables—indicating the indirect effects that education and gender bias in education have on economic growth. Path analysis can then be used to determine the total effect, defined as: Total effect equals direct effect plus indirect effects.[20] For example, the total effect of the initial female-male ratio of schooling (*RED60*) on growth is:

$$\beta_6 + (\beta_{13} * \beta_1) + (\beta_{18} * \beta_2) + (\beta_{18} * \beta_9 * \beta_1) + (\beta_{23} * \beta_3) + (\beta_{23} * \beta_{10} * \beta_1).$$

The first term is the direct effect, the second term is the indirect effect through investment, the third term is the indirect effect through population growth, the fourth term is the indirect effect through population growth and investment, the fifth term is the indirect effect through labor force growth, and the sixth term is the indirect effect through labor force growth and investment.

Equation (5) is a "reduced form" regression that omits investment, population growth, and labor force growth. Thus it should measure the total effect of gender bias in education directly.

It is important to briefly discuss how human capital and gender bias in human capital are modeled here.[21] Instead of including variables for male and female human capital achievement—which have a correlation coefficient of 0.96 between initial levels and 0.69 between growth in education—a different approach was used to avoid the multicollinearity problem. The regressions include a variable that measures overall human capital (*ED60* and *GED*) as well as one that measures the female–male ratio of human capital (*RED60* and *RGED*). The first tries to capture the effect on growth of overall human capital, and the second measures whether a country with smaller gender differences in education would grow faster than a country with identical average human capital but greater inequality in its distribution. This approach generates much smaller correlation coefficients and should make it easier to identify the different effects.[22]

20. An important assumption in interpreting such a path analysis is that one has good reason to believe that the causality in the indirect regressions runs from gender bias in education to investment, population growth, and labor force growth. That seems to be true in these three cases, which is why only these three indirect effects are considered here.

21 An important issue is whether a stock or flow measure of human capital is more appropriate for such an estimation. Because the goal here is to model economic growth, it appears appropriate to focus on flow measures, such as physical and human capital investments, and thus use the change in years of schooling as a proxy for investments in human capital At the same time, it may be that due to externalities and complementarities between factors of production, the stock of human capital increases economic growth because it makes physical capital more productive. Thus it is useful to include both a stock and a flow measure, as proposed for the regressions

22. The correlation coefficient between initial levels (i.e., *ED60* and *RED60*) is now reduced to 0.61 or 0.68 and between growth rates of education (*GED* and *RGED*) it is now –0.29 or –0 1, depending on whether male or average education is used to proxy for average education.

Two assumptions can be made about the relationship between male education and the gender gap in education. The first implicitly assumes that gender inequality in education could be reduced without reducing male education levels. In this case, male years of schooling can be used to proxy for average human capital (*ED60* and *GED*). Given that increases in female education might at least partially come at the expense of male education, this assumption provides an upper bound of the measured effect of gender inequality on growth. In the second assumption any increase in female schooling, other things being equal, would lead to a commensurate decrease in male schooling. In this case the average of male and female years of schooling is used for *ED60* and *GED* and thus provides a lower-level estimate of the measured growth effect of gender inequality in education. The true effect likely falls between these two estimates, probably closer to the first than the second. Both specifications are reported on later in the discussion.[23]

A second issue relates to possible simultaneity problems. Because the estimation relies not only on initial levels of and gender differences in human capital (*ED60* and *RED60*) but also on growth in education attainment and the female–male ratio of that growth (*GED* and *RGED*), it could theoretically be that causality runs from growth to schooling (and reduced gender bias in schooling; see Dollar and Gatti 1999)—not the other way around, as suggested here.

This issue is addressed using three procedures. First, total years of schooling of the adult population is a stock measure of education based on past investments in education. Growth in total years of schooling of the adult population between 1960 and 1990 (*GED*) is largely based on education investments between 1940 and 1975. Thus it is unlikely (but not impossible) that these investments were mainly driven by high economic growth between 1960 and 1992. In addition, some regressions use a stock measure of education among people age 25 and older in 1970—a measure that cannot be heavily influenced by growth after 1960.[24]

Second, instrumental variable techniques are used to address this issue. In particular, education spending (as a share of GDP), initial fertility levels, and changes in fertility rates are used as instruments for changes in education attainment and in the female–male ratio of those changes to explicitly control for possible simultaneity. These instruments pass standard relevance and overidentification tests and so appear to be plausible candidates.[25]

23. Another way to model gender inequality in education would be to have a variable for male education and a second one for the difference between female and male education The results of that approach are very similar to the results when using ratios

24. It is not possible for this measure of human capital to have been greatly influenced by economic growth because the investments (especially in primary and secondary education) needed to educate people who were 25 and older in 1970 were completed between 1930 and 1960 and so cannot have been influenced by economic growth since 1960.

25. The instruments are much better at predicting the female–male ratio of growth in education than male growth in education (where they have only a weakly significant impact) To see whether these instruments are appropriate in the sense that they affect the dependent variable only through the variable they instrument and not directly, the overidentificiation restriction test was run. To do so, the residuals from the second-stage regressions were saved and regressed on the exogenous variables.

Third, the model is reestimated using a panel dataset that divides the dependent and independent variables into three periods (1960–70, 1970–80, and 1980–90). These panel regressions use only initial values of education and gender gaps in initial education as explanatory variables.[26]

These models are estimated for 109 industrial and developing economies. In addition, the same analysis is performed for a sample that includes only developing countries and for a sample that includes only African countries to see whether the relationships differ.[27]

IV. Descriptive Statistics

This section presents regional data on growth and its most important determinants as well as on gender inequality in education and employment. These data form the background for the multivariate analysis in the next section.

Between 1960 and 1992 annual growth in per capita income was slowest in Sub-Saharan Africa, averaging just 0.9 percent (table 1). This was about 40 percent of the world average and 3.3 percentage points less than in East Asia and the Pacific. Latin America also experienced slow growth, followed closely by South Asia. Similarly, average investment rates were low in Sub-Saharan Africa but slightly higher than in South Asia.

Moreover, Sub-Saharan Africa, South Asia, and the Middle East and North Africa suffered from several disadvantages in initial conditions. Between 1960 and 1992 these three regions had the world's highest population growth. Moreover, in 1960 the average woman (age 15 or older) in Sub-Saharan Africa and the Middle East and North Africa had a dismal 1.09 years of schooling; in South Asia female education attainment was even worse. Gender inequality in education was also high in all three regions, with women in Sub-Saharan Africa having barely half the schooling of men and women in South Asia and the Middle East and North Africa having only about a third.

Particularly worrisome is that between 1960 and 1990 education expansion in Sub-Saharan Africa was the lowest of all regions. Moreover, the region's female–male ratio in that growth was 0.89, meaning that females experienced

---

the product of N (number of observations) and the R^2 is distributed as a chi-square, and the null hypothesis is that the exogenous variables have no impact on the residuals The relevant statistic is 0.49—clearly unable to reject the hypothesis and so confirming that the instruments are appropriate.

26. Specification tests indicate that the best panel regression specification is to use regional and decade dummies but no country-specific fixed effects or random effects. The Breusch-Pagan test for random effects (P-value of 0.2083) suggests that there is no reason to believe that there are random or fixed effects. Moreover, a fixed regression specification in a panel with only three time-series observations often leads to inconsistent results. The specification with regional and decade dummies has the highest explanatory power and easily passes the Ramsey Reset test for omitted variables. The fixed and random effects specifications yield qualitatively the same results.

27. All standard errors are adjusted for heteroscedasticity, and the Ramsey Reset test is used to investigate the presence of omitted variables.

TABLE 1. Economic and Demographic Indicators by Region, 1960–92

Region	Growth (g)	*Inv*	OPEN	*Popgro*	LFG	Initial income	Final income	FED60	MED60	RED60	RED70	MGED	RGED
Average	2.1	18.5	62.5	1.9	2.2	2,422	5,166	3.22	4.01	0.69	0.67	0.07	1.01
East Asia and Pacific	4.2	21.0	98.2	2.2	2.7	1,283	5,859	2.79	4.67	0.59	0.56	0.09	1.44
Eastern Europe and Central Asia	3.2	33.3	54.1	0.4	0.5	2,675	5,308	6.16	7.38	0.83	0.83	0.08	1.26
Latin America and Caribbean	1.3	16.3	57.4	2.1	2.6	2,302	3,471	3.27	3.68	0.89	0.83	0.06	1.11
Middle East and North Africa	2.3	16.0	56.5	2.9	3.1	1,776	3,518	1.09	2.09	0.37	0.34	0.12	0.82
South Asia	1.7	8.9	30.4	2.3	2.7	760	1,306	0.87	1.84	0.33	0.31	0 08	0.77
Sub-Saharan Africa	0.9	10.8	65.1	2.7	2.8	855	1,251	1.09	1.85	0.51	0.49	0.06	0.89
OECD	2.9	27.0	59.8	0.7	0.9	5,583	12,870	6.48	6.86	0.93	0.89	0.07	0.90

Note: Data cover 109 countries for which all covariates are available. Regional data are unweighted averages for the countries in each region. For example, female–male ratios for education are calculated as the regional averages of the ratios in each country (not by dividing the regional average for female education by the regional average for male education). Thus dividing FED60 by MED60 for a particular region will not equal RED60 for that region. Variables are defined as follows: *Growth* (g) average annual compounded rate of per capita income growth (percent) in 1960–92; *Inv*: average annual rate of investment as a share of GDP (percent) in 1960–92; OPEN average annual ratio of exports and imports as a share of GDP (percent) in 1960–92; *Popgro:* average annual compounded rate of population growth (percent) in 1960–92; LFG: average annual compounded rate of growth in the labor force (ages 15–64) population (percent) in 1970–92; Initial income: average per capita income in 1960 (or first year available), in constant 1985 U.S. dollars using the chain index; Final income: average per capita income in 1992 (or last year available), in constant 1985 U.S. dollars using the chain index; FED60: average years of schooling among women age 15 and older in 1960, MED60: average years of schooling among men age 15 and older in 1960 (used for average education variable ED60 in most regressions below); RED60: female–male ratio of total years of schooling among people age 15 and older in 1960; RED70: female–male ratio of total years of schooling among people age 25 and older in 1970; MGED. average annual absolute growth in years of schooling among men age 15 and older in 1960–90 (used for average education variable GED in most regressions); and RGED: female–male ratio of the average annual absolute growth in total years of schooling among people age 15 and older in 1960–90.

Source: See text.

slower expansion in education achievement than males. South Asia and the Middle East and North Africa were equally poor performers. Although male education levels expanded much faster than in Sub-Saharan Africa, females lagged even further behind in education expansion. Again, the contrast with East Asia and the Pacific is striking. There, female years of schooling expanded three times faster than in Sub-Saharan Africa and female education expansion outpaced male expansion by 44 percent.

Thus Sub-Saharan Africa, South Asia, and the Middle East and North Africa suffered from the worst initial conditions for female education and had the worst record of improvements between 1960 and 1990. In contrast, East Asia and the Pacific started out with somewhat better conditions for women's education. But more important, women were able to improve their education levels much faster than men, rapidly closing the initial gaps.[28]

So, if gender inequality in education affects economic growth, Sub-Saharan Africa, South Asia, and the Middle East and North Africa should have suffered the most damage, having experienced the highest gender inequality in initial education and in its expansion. Meanwhile, East Asia and the Pacific and Eastern Europe and Central Asia should have benefited from lower initial and rapidly falling inequality in education.

V. Econometric Analysis

The basic regressions for equations (1–5) are shown in table 2. All the regressions have a high explanatory power and perform well on specification tests. Regression 1 confirms a number of known findings on conditional convergence (*LNINC1960*), the importance for growth of investment and openness (*Inv*, *OPEN*), the importance of initial levels of human capital (*ED60*) and growth in human capital (*GED*), the negative impact of population growth (*Popgro*), and the positive impact of labor force growth (*LFG*; see ADB 1997, Barro 1991, Bloom and Williamson 1998, and Mankiw, Roemer, and Weil 1992). The coefficients for these variables are within the ranges observed in other studies. Some of the dummy variables for different regions are significant—especially for Sub-Saharan Africa and Latin America—suggesting that the growth regression is not picking up all the effects that account for slower growth in these regions (see also Barro 1991).

More interesting for the purposes of this article is the finding that both the initial female–male ratio of schooling (*RED60*) and the female–male ratio of expansion in schooling (*RGED*) have a significant positive impact on economic

28. Eastern Europe is also notable for its small gender gaps in education and employment (see Klasen 1993). In addition, note that Sub-Saharan Africa and South Asia had the world's lowest incomes in 1960, which should have helped boost growth because trade and factor flows should have promoted convergence in income levels. Sub-Saharan economies exhibited average levels of openness, whereas South Asian economies were more closed. But this comparison is slightly deceptive, because one would have expected the many small Sub-Saharan economies to have above-average levels of openness due to their small domestic markets. Relative to East Asia, Sub-Saharan economies were much less open.

TABLE 2. Gender Inequality in Education and Economic Growth

Dependent variable	(1) Growth	(2) *Inv*	(3) *Popgro*	(4) *LFG*	(5) Growth	(6) Growth	(7) Growth
Constant	6.33***	3.23	3.74***	4.16***	7.50***	6.82***	7.89***
	(3.4)	(0.5)	(4.5)	(4.6)	(4.7)	(3.8)	(5.0)
LNINC60	-1.13***	-0.10	-0.13	-0.18*	-1.21***	-1 16***	-1.24***
	(-5.0)	(-0.1)	(-1.1)	(-1.4)	(-5.3)	(-5.2)	(-5.4)
Popgro	-0.55*	-0.81	—	—	—	-0.47	—
	(-1.4)	(-0.8)	—	—	—	(-1.1)	—
LFG	0 62*	2.60**	—	—	—	0.53	—
	(1.5)	(2.2)	—	—	—	(1.2)	—
OPEN	0.007**	0.026**	—	—	0.009***	0.007**	0.010***
	(1.9)	(1.9)	—	—	(2.6)	(2.1)	(2.7)
Inv	0.056**	—	—	—	—	0.055**	—
	(1.7)	—	—	—	—	(1.7)	—
ED60	0.19**	0.64**	-0.03	-0.02	0.23***	0.18**	0.21**
	(2.3)	(1.9)	(-0.5)	(-0.6)	(2.5)	(2.1)	(2.3)
GED	12.61***	14.10	-0.86	0.78	14.38***	14.93***	17.47***
	(3.8)	(0.9)	(-0.6)	(0.5)	(3.7)	(4.1)	(4.2)
RED60	0.90*	7.44***	-0.29	0.19	1.64**	0.78	1.46**
	(1.3)	(2.7)	(-1.1)	(0.8)	(2.3)	(1.1)	(2.0)
RGED	0.69***	1.28	-0.22***	-0.13*	0.75***	0.51**	0.55***
	(3.0)	(1.2)	(-2.4)	(-1.4)	(2.8)	(2.2)	(2.2)
Eastern Europe and	-0.77	14.82***	-1.56***	-2.12***	-0.57	-0.74	-0.48
Central Asia	(-0.9)	(5.5)	(-7.1)	(-10.4)	(-0.8)	(-0.9)	(-0.7)
Latin America and	-1.31**	-4.12**	-0.01	-0.12	-1.58**	-1.30**	-1.56**
Caribbean	(-1.8)	(-2 1)	(-0.1)	(-0.5)	(-2.2)	(-1.9)	(-2.3)
Middle East and	-0.15	-0.44	0.50**	0.25	-0.25	-0.10	-0.19
North Africa	(-0.2)	(-0.2)	(2.2)	(1.1)	(-0.4)	(0.2)	(-0.3)
South Asia	-0.46	-5.35**	-0.20	-0.25	-0.78	-0.42	-0.72
	(-0.7)	(-2.4)	(-0.9)	(-1.1)	(-1.1)	(-0.6)	(-1.1)
Sub-Saharan	-1.42**	-5.53***	0.27	-0.10	-1.95***	-1.38**	-1.85***
Africa	(-2.1)	(-2.8)	(1.1)	(-0 5)	(-3.0)	(-2.0)	(-3.0)
OECD	0.49	7 89***	-1.23***	-1.64***	0.46	0.59	0.63
	(0.7)	(3.6)	(-6.1)	(7.5)	(0.6)	(0.9)	(0.9)
Adjusted R^2	0.61	0.74	0.65	0.71	0.57	0.61	0.59
Omitted variables test	Passed	Passed	Passed	Passed	Passed	Passed	Passed
N	109	109	109	109	109	109	109

— Not available.

*Significance at the 90 percent level (one-tailed test).

**Significance at the 95 percent level (one-tailed test).

***Significance at the 99 percent level (one-tailed test).

Note: Numbers in parentheses are heteroscedasticity-adjusted *t*-ratios. *LNINC60* refers to the log of income per capita in 1960 adjusted for purchasing power parity. In regressions 1–5 *ED60* refers to average years of schooling among men age 15 and older in 1960, and *GED* refers to average annual absolute growth in years of schooling among men age 15 and older in 1960–90 In regressions 6 and 7 *ED60* and *GED* refer, respectively, to average years of, and average annual absolute growth in schooling for both men and women. Other variables are explained in table 1. The Ramsey Reset test is used to test for omitted variables. In regressions 1 and 6 the Ramsey Reset test is passed only when powers of the right-hand-side variables are considered (not when powers of the fitted values for the dependent variable are considered, as in all the other regressions). East Asia and the Pacific is omitted.

Source: Author's calculations.

growth. Because the regressions control for investment as well as population and labor force growth, these results support the claim that gender inequality in education reduces average human capital. They also support the direct externality effect of increased female education. The magnitude of the coefficient is plausible. An increase from 0.5 to 1.0 in the female–male ratio of growth in schooling would raise the annual growth rate by about 0.4 percentage points, similar to the predictions made earlier.[29]

Regression 2 shows the determinants of investments and finds that higher investment rates are related to higher labor force growth, greater openness, and higher human capital. In addition, reduced gender inequality in education appears to lead to higher investment rates, confirming the indirect links between gender inequality in education, investment, and economic growth postulated. In particular, the regression confirms the impact that the quality of human capital has on investment rates. Moreover, it is consistent with the claim that lower gender inequality, combined with wage discrimination boost investment because it makes investing in female-intensive industries especially appealing. The effect through the investment rate seems particularly large for initial gender inequality in education.

Regressions 3 and 4 show that gender inequality in education has the expected impact on population growth and labor force growth, so the indirect link between gender inequality in education and economic growth through these two factors is also present.

Regression 5 shows the "reduced form" estimate of the impact of gender inequality in education. Comparisons between regressions 1 and 5 indicate that the indirect effects of gender inequality in education are large because the size (and significance) of both coefficients—particularly the one related to initial gender inequality—have increased considerably. This finding suggests that initial gender inequality mainly affects growth indirectly, particularly through its effect on investment rates.

In table 3 the results from regressions 1–5 are used to determine the extent to which growth in Sub-Saharan Africa, South Asia, and the Middle East and North Africa has lagged behind growth in East Asia and the Pacific due to initial gender bias in education and gender bias in the growth of education. As this specification used male education to proxy for average human capital, it is implicitly assuming that the female education could have been increased without chang-

29. Empirically, gender inequality in education also appears to be related to the health of the population. When the under-age-5 mortality rate or life expectancy in 1960 are included in the regressions (not shown here), the direct effects on growth of gender inequality in education become smaller (but remain sizable), and the coefficients on child mortality and longevity are in the right direction but not significant. As shown in Klasen (1999), lower gender inequality in education reduces child mortality—suggesting that lower gender inequality also promotes economic growth through its effect on lower child mortality and thus a healthier population. This finding also suggests that the inclusion of initial life expectancy in growth regressions (as in ADB 1997) picks up part of this effect, which is really due to gender inequality in education.

TABLE 3. Gender Inequality in Education and Its Effects on Growth Differences between East Asia and the Pacific and Other Regions
(percent)

Indicator	Upper-bound estimate of growth difference between East Asia and the Pacific and:			Lower-bound estimate of growth difference between East Asia and the Pacific and:		
	Sub-Saharan Africa	South Asia	Middle East and North Africa	Sub-Saharan Africa	South Asia	Middle East and North Africa
Total annual growth difference	3.3	2.5	1.9	3.3	2.5	1.9
Accounted for by						
Direct effect of gender inequality in education (1)	0.45 (0.08, 0.37)	0.69 (0.23, 0.46)	0.63 (0.20, 0.43)	0.34 (0.06, 0.28)	0.54 (0.20, 0.34)	0.49 (0.17, 0.31)
Indirect effect through investment (2)	0.07	0.16	0.14	0.07	0.14	0.13
Indirect effect through population growth (2, 3)	0.09	0.13	0.12	0.07	0.11	0.10
Indirect effect through labor force growth (2, 4)	–0.04	–0.03	–0.03	–0.04	–0.02	–0.02
Total direct and indirect effect (sum of above)	0.56 (0.13, 0.43)	0.95 (0.43, 0.53)	0.86 (0.37, 0.49)	0.44 (0.12, 0.32)	0.77 (0.38, 0.39)	0.69 (0 33, 0.36)
Total effect using reduced form regression (5)	0.54 (0.13, 0 41)	0.92 (0.43, 0.50)	0.83 (0.37, 0.46)	0.42 (0.12, 0.30)	0.75 (0.38, 0.37)	0 67 (0 33, 0.34)

Note. The numbers in parentheses in the first column refer to the regressions on which the estimates are based. The first number in parentheses in the other columns refers to the difference in growth accounted for by different initial levels of gender inequality in education (based on the variable *RED60*) and the second number refers to the difference accounted for by the gender gap in growth of schooling (based on the variable *RGED*). Barring rounding errors, the two numbers sum to the number reported before the parentheses—the combined effect.

Source: Author's calculations.

ing male education levels, the table present upper-bound estimates of the effects (see later discussion for a discussion of lower-bound estimates, which are shown in the right-hand panel of table 3). The data indicate the combined effects (and in parentheses, disaggregated effects) of initial gender gaps in education and gender gaps in the growth of education. Using just the direct effect (regression 1), 0.45 of the 3.3-percentage-point annual difference in growth between Sub-Saharan Africa and East Asia can be accounted for by differences in gender inequality in education, with most due to differences in gender bias in the growth of education.[30]

The comparison between South Asia and East Asia is even more striking. Here 0.69 of the regions' 2.5-percentage-point annual difference in economic growth can be accounted for by differences in gender inequality in education. About two-thirds is explained by differences in gender bias in the growth of education and one-third by differences in gender inequality in 1960. Similarly, gender inequality in education appears to have slowed growth in the Middle East and North Africa by amounts similar to those in South Asia.

Indirect effects also account for some of the differences in economic growth between South Asia (and Sub-Saharan Africa) and East Asia. Through its effect on investment, gender inequality in education accounts for another 0.16 percentage point of the growth difference between South Asia and East Asia (and for 0.07 percentage point of the difference between Sub-Saharan Africa and East Asia). Gender inequality in education also accounts for 0.13 percentage point of the growth difference between South Asia and East Asia through its effect on population growth (0.09 percentage point of the difference Africa-East Asia). These indirect effects—especially those operating through population and labor force growth rates—are somewhat smaller than expected and considerably smaller than the direct effect of gender inequality.

The total direct and indirect effects of gender inequality in education account for 0.95 percentage point of the difference in economic growth between South Asia and East Asia, 0.56 percentage point of the difference between Sub-Saharan Africa and East Asia, and 0.85 percentage point of the difference between the Middle East and North Africa and East Asia. Using the reduced form regression (regression 5) yields nearly identical estimates of the total size of the effects as well as their impact on the growth differences between regions (see table 3).

Thus gender inequality in education appears to have a sizable effect on economic growth. It is worth emphasizing that the results in table 3 do not take into account differences in average human capital between regions, just gender inequality in education. The regressions show that differences in average human

30. Through the process of conditional convergence, Sub-Saharan Africa could have been expected to grow faster than East Asia. If this factor is taken into account, the growth difference to be explained increases to 3.9 percentage points

capital also matter a lot and can account for an additional share of the growth differences between regions.

As noted, it is important to determine the robustness of the results presented. First, as already mentioned, the estimates on gender inequality in education present an upper-bound estimate because they assume that increases in female education could have been achieved without any reduction in male education. Columns 6 and 7 in table 2 present the direct effect and reduced form regressions for the lower-bound estimate,[31] using the average level of human capital rather than the male level of human capital used previously.

As expected, the coefficients for the gender gap in schooling are smaller. But they are still sizable and, in most cases, significant. Calculations of the growth differences accounted for by this measure of gender inequality in education show only small differences from the previous ones (see table 3). Here the total effect of gender inequality in education accounts for 0.77 percentage point of the growth difference between South Asia and East Asia, 0.44 percentage point of the difference between Sub-Saharan Africa and East Asia, and 0.69 percentage point of the difference between the Middle East and North Africa and East Asia.

Second, there is a need to consider possible simultaneity issues. In particular, is it possible that growth led to increases in the female–male ratio of schooling rather than the other way around? This issue is first addressed by replacing the two education gap variables with just one that measures the female–male ratio of schooling of adults above 25 in 1970 (as noted, this statistic could hardly have been affected by economic growth after 1960). In regressions not shown here, this measure has a significant effect on economic growth—supporting the notion that causality runs from gender bias in education to growth, not vice versa.

Simultaneity is also addressed in regressions 8 and 9 in table 4, which are based on a panel analysis in which the dependent and independent variables are divided into three periods. Because only initial levels of schooling and initial female–male ratios of schooling are included in the regression, it avoids the simultaneity issue inherent in the education growth variables. The results are very similar to the cross-section results, which is reassuring because findings from cross-country regressions often change in a panel setting. In fact, even the magnitude of the effects is roughly similar to that in the cross-section regression.

Table 5 estimates the impact of gender inequality in education on differences in growth between regions using the panel regressions. In the 1980s initial gender inequality in education accounted for about 0.3–0.5 percentage point of the growth differences between East Asia and Sub-Saharan Africa, South Asia, and the Middle East and North Africa. These differences are smaller than in the cross-section regression, which is to be expected because the analysis no longer considers the impact of further improvements in the gender gap in education that may have occurred after 1980.

31. Here and in the following, the intervening regressions of the indirect effects are available on request.

TABLE 4. Gender Inequality in Education and Economic Growth: Further Specifications

Dependent variable	(8) Growth	(9) Growth	(10) Growth	(11) Growth	(12) Growth	(13) Growth	(14) Growth
Constant	6.33***	3.23	3.74***	4.16***	7.50***	6.82***	7.89***
Constant	7.92***	7.68***	4.34	4.42**	5.90***	1.29	3.91*
	(3.5)	(3.5)	(1.1)	(2.2)	(3.6)	(0.3)	(1.6)
LNINC60	-1.15***	-1.12***	-1.30*	-0.96***	-1.04***	-0.80**	-1.03**
	(-3.5)	(-3.5)	(-1.6)	(-3.85)	(-4.2)	(-1.9)	(2.5)
Popgro	-1.02***	—	—	-0.45	-1.19	—	—
	(-3.1)	—	—	(-1.1)	(-1.2)	—	—
LFG	0.61***	—	—	0.70*	—	1.66*	—
	(2.4)	—	—	(1.5)	—	(1.3)	—
OPEN	0.007**	0.012***	0.001	0.008**	0.010***	0.015**	0.012**
	(2.1)	(3.7)	(0.1)	(2.2)	(2.6)	(2.0)	(1.8)
Inv	0.111***	—	—	0.035	—	0.019	—
	(4.7)	—	—	(0.4)	—	(-0.4)	—
ED60	0.08	0.19**	0.29	0.31***	0.33***	0.06	0.19
	(1.0)	(1.9)	(0.8)	(3.0)	(3.2)	(0.2)	(0.7)
GED	—	—	33.96	14.1***	15.39***	11.18**	11.78**
	—	—	(1.2)	(3.9)	(3.2)	(2.2)	(2.4)
RED60	1.33**	2.12***	1.24	1.14*	1.71**	2.59**	3.29***
	(1.65)	(2.6)	(0.8)	(1.5)	(2.2)	(2.1)	(3.0)
RGED	—	—	2.64**	0.53**	0.52**	0.62*	0.36
	—	—	(1.7)	(2.2)	(2.0)	(1.5)	(1.1)
Eastern Europe and	-1.83**	-0.70	0.26	-0.58	-0.98*	—	—
Central Asia	(-2.3)	(-0.9)	(0.3)	(-0.6)	(-1.3)	—	—
Latin America and	-1.87***	-2.29***	-0.50	-1.39**	-1.62**	—	—
Caribbean	(-3.0)	(-3.5)	(-0.4)	(-1.8)	(2.2)	—	—
Middle East and	0.22	0.08	0.14	-0.12	-0.16	—	—
North Africa	(0.3)	(0.1)	(0.2)	(-0.2)	(-0.2)	—	—
South Asia	-1.01*	-1.78**	0.40	-0.24	-0.48	—	—
	(-1.4)	(-2.3)	(0.4)	(-0.35)	(-0.7)	—	—
Sub-Saharan Africa	1.74***	-2.76***	-0.35	-1.26**	-1.64***	—	—
	(-2.9)	(-4.5)	(-0.3)	(-1.7)	(2.6)	—	—
OECD	-1.01*	-0.34	1.73**	—	—	—	—
	(-1.6)	(-0.5)	(1.9)	—	—	—	—
1960s	2.60***	2.58***	—	—	—	—	—
	(7.9)	(7.9)	—	—	—	—	—
1970s	1.74***	2.10***	—	—	—	—	—
	(5.9)	(6.9)	—	—	—	—	—
Adjusted R^2	0.47	0.40	0.19	0.60	0.58	0.55	0.56
Omitted variables test	Passed	Passed	Not applicable	Passed	Passed	Passed	Passed
N	295	295	96	86	86	27	27

— Not available.

*Significance at the 90 percent level (one-tailed test).

**Significance at the 95 percent level(one-tailed test).

***Significance at the 99 percent level (one-tailed test).

Note: Numbers in parentheses are heteroscedasticity-adjusted *t*-ratios. *LNINC60* refers to the log of income per capita in 1960 adjusted for purchasing power parity. Other variables are explained in table 1 This table shows only the direct effect and reduced form regressions. The intervening regressions of investment, population growth, and labor force growth are available on request. Regressions 8 and 9 are based on a panel regression with three observations per country (1960s, 1970s, and 1980s). In these regressions *ED60* and *RED60* refer to the level and female-male ratio of years of schooling of people age 25 and older at the start of each decade. Regression 10 is the second stage of a two-stage least squares regression. The instruments used for *GED* and *RGED* are government spending on education (as a share of GDP), the fertility rate in 1960, and the change in the fertility rate between 1960 and 1990. Regressions 11 and 12 restrict the sample to developing economies, and regressions 13 and 14 to countries in Sub-Saharan Africa. The Ramsey Reset test is used to test for omitted variables. East Asia and the Pacific is omitted.

Source: Author's calculations.

TABLE 5. Gender Inequality in Education and Its Effects on Growth Differences between East Asia and the Pacific and Other Regions: Panel-Based Estimates
(percent)

Indicator	Estimate of growth difference between East Asia and the Pacific and: Sub-Saharan Africa	South Asia	Middle East and North Africa
Growth difference, 1960s	2.13	2.49	1.26
Growth difference, 1970s	3.82	3.77	0.90
Growth difference, 1980s	3.04	1.54	2.97
Difference in female–male ratio of years of schooling, 1960	0.06	0.18	0.11
Difference in female–male ratio of years of schooling, 1970	0.08	0.23	0.16
Difference in female–male ratio of years of schooling, 1980	0.17	0.21	0.16
Difference accounted for by gender inequality in education			
Direct effect (10)			
1960s	0.09	0.24	0.15
1970s	0.11	0.30	0.22
1980s	0.23	0.28	0.21
Total effect using reduced form regression (11)			
1960s	0.14	0.38	0.25
1970s	0.17	0.49	0.34
1980s	0.36	0.45	0.34

Note. Estimates are based on regressions 10 (direct effect) and 11 (total effect using reduced form regression), which are panel regressions with one observation per decade. These regressions measure only the impact of the female–male ratio of years of schooling among people age 25 and older at the beginning of each decade (and thus not the impact of the change in education achievements throughout a decade).
Source: Author's calculations

The panel regressions also show an interesting temporal pattern in the impact that gender inequality in education has on economic growth. In 1960 East Asia did not exhibit much lower gender bias in education than did other regions, and growth differences between it and other regions were comparatively small. By 1970 and especially by 1980, the gender gap in East Asia was much lower than in other regions—and it was precisely during that period when growth differences soared between East Asia and South Asia, Sub-Saharan Africa, and the Middle East and North Africa. These findings suggest that smaller gender gaps in education played a significant role in higher growth.

To approach the simultaneity issue in another way, regression 10 presents a two-stage least squares regression in which growth in average education and in the female–male ratio of education are replaced by their predicted values, using as instruments government spending on education (as a share of GDP), the fertility rate in 1960, and the change in the fertility rate between 1960 and 1990 (see

table 4). The female–male ratio of growth of education still has a significant impact and is much larger in magnitude. This finding lends further support to the contention that causality runs from gender bias in education to economic growth, not the reverse.

Further analyses investigate whether the relationship between gender inequality and growth differs depending on which countries are included in the regression. Although limiting the sample to more homogeneous groups of countries makes it possible to see whether effects differ by region, it eliminates some of the important cross-sectional variation needed to estimate those effects, which may lead to less precise results. Limiting the sample to 85 developing economies (regressions 11 and 12 in table 4) has little effect on the results. Initial gender inequality has a slightly larger effect than in the full sample, and gender bias in the growth of education has a smaller effect. Using these regressions to account for growth differences suggests that gender inequality in education accounts for 0.44 percentage point of the growth difference between Sub-Saharan Africa and East Asia and 0.81 percentage point of the difference between South Asia and East Asia.

Thus gender inequality in education has as much effect on growth in developing economies as in industrial economies. This finding differs from that in Dollar and Gatti (1999), who find that gender inequality in education significantly affects growth only in countries with high female education (see later discussion).

Limiting the sample to Sub-Saharan Africa (27 observations) produces some interesting results, as shown in regressions 13 and 14 in table 4. Initial gender inequality in education has a much larger effect than in the full sample, whereas gender bias in the growth of education has a similar effect. In sum, gender inequality in education appears to matter more in Sub-Saharan Africa than elsewhere, despite the agrarian nature of most African economies. This finding suggests that, given women's important role in agriculture in Sub-Saharan Africa (and elsewhere), their poor human capital appears to be a particularly significant constraint to economic growth. For example, about 30 percent of the large growth differences between Botswana (which had low initial education gaps and closed them rapidly) and Ghana and Niger (which still have gender inequality in education) can be accounted for by differences in gender inequality in education.

VI. Reconciling These Results with Those of Other Studies

This article has found that gender inequality in initial education levels and in the expansion of education significantly reduces economic growth. These effects are visible in a large sample of industrial and developing economies but are just as evident in a sample only of developing economies—and are especially apparent in Sub-Saharan Africa. The results are robust to different specifications of education variables and controls for possible simultaneity. How do these results compare with those of previous studies?

The results are qualitatively consistent with those of Hill and King (1995) and Knowles, Lorgelly, and Owen (2002), who achieve similar results based on regres-

sions estimating GDP levels. The size of the coefficients are not directly comparable because level regressions (which measure differences in steady states) have different coefficients than growth regressions (which measure transitions to steady states).[32]

The results here are the opposite of Barro and colleauges, who find that female education has a negative effect on growth. I can reproduce Barro's results when I use a specification similar to his, but those results disappear—and indeed, become reversed—once the multicollinearity problem is addressed and regional dummy variables are added. Clearly, his results are largely driven by these two problems.

The findings in this article also contrast somewhat with those of Dollar and Gatti (1999).[33] Although both sets of findings emphasize the negative effect on growth of gender inequality, Dollar and Gatti find that this effect is concentrated among industrial countries. But their study differs from this one in two important ways. First, they use a much shorter period (1975–90) and a three-part panel in that period. Second, they use a different education variable—one that measures the share of adults who have *exactly* achieved some secondary education. This variable appears problematic because it ignores people who have completed secondary education or even have some tertiary education.

If I use their specification and time period but replace their variable with a more plausible one—namely, the share of adults that have achieved *at least* some secondary education—the finding that the gender gap in education hurts growth only in industrial countries disappears. In fact, although the coefficients on gender inequality in this specification are generally not significant, gender inequality in education has a stronger negative effect on growth in developing countries, which is closer to the findings presented here.

If I replace their variable with one more like the one I used—namely, average years of primary and secondary schooling among adults—the results are much closer to my findings. In particular, female education has a strong and significant positive effect on growth in developing economies, and the effect is weaker or even nonexistent in industrial countries. Thus the choice of education variable appears to matter. As argued, it is more plausible to use a variable that measures the average total years of schooling or the share of the population with at least some secondary education.

The choice of time period also appears to matter. When the panel dataset and specification used here are restricted to 1970–90, the direct effect on growth of gender bias in education becomes smaller and falls just below conventional levels of significance in industrial and developing economies separately and when com-

32. Knowles, Lorgelly, and Owen (2002) also present a growth regression, and in a similar specification their coefficient is quite similar to the result found here.

33. I am grateful to Roberta Gatti for providing the database underlying their results. Although I was not able to reproduce their results exactly—because of differences in the number of observations and questions about the use of constant and regional dummy variables—I was able to do so qualitatively and so reconcile them with mine.

bined. But when the reduced regressions are run, the effects remain strong and significant and larger in developing economies. If the time period is further restricted to 1980–90, the effects are even weaker and less significant, although the qualitative story remains the same.

The years of 1975–90 and, more extremely, 1980–90, were exceptional periods for most developing economies. Most suffered negative external shocks, with serious balance of payments and debt crises, forcing them to adopt structural adjustment programs. In such unusual times it is no surprise that human capital variables that pay off only in the long term do not show significant effects.

Thus, the findings here are consistent with earlier studies or exhibit variations that can be traced to different time periods, education variables, specifications, and underlying data. At the same time, the results use the most comprehensive time period, the most plausible education variable, and the most comprehensive robustness checks and show that the effects are found in different subsamples.

VII. Conclusion

This article has examined the extent to which gender inequality in education reduces economic growth. Several findings are especially important. First, gender inequality in education undermines economic growth directly by lowering average human capital and indirectly through its impact on investment and population growth. The effects are sizable. If Sub-Saharan Africa, South Asia, and the Middle East and North Africa had started with more balanced education achievements in 1960 and done more to promote gender-balanced education growth, their annual economic growth rates could have been up to 0.9 percentage point faster.

Second, these effects do not appear to be related to simultaneity issues. Several specifications and the use of instrumental variable estimation show that gender inequality in education has a persistent effect on economic growth. Third, these effects appear to be stronger in Sub-Saharan Africa—suggesting that efforts to promote female education have a higher payoff there than elsewhere.

Thus promoting gender equity in education may be among the few "win–win" development strategies. It advances economic prosperity and efficiency, promotes other essential human development goals (such as lower mortality and fertility), and is intrinsically valuable as well.

Although these results appear to be robust, plausible, and in line with expected theoretical effects, the findings presented here, as with all empirical growth regressions, show associations but cannot prove causality. It is possible that the findings are partly due to the omission of some variable not considered, that measurement error affects the results, or that the model is misspecified in other ways. Further investigations of this nature—as well as complementary analyses using micro-data—are needed to corroborate the findings and the mechanisms underlying them.

References

Alderman, H., J. Behrman, D. Ross, and R. Sabot. 1996. "Decomposing the Gender Gap in Cognitive Skills in a Poor Rural Economy." *Journal of Human Resources* 31(1):229–54.

Alderman, H., J. Behrman, S. Khan, D. Ross, and R. Sabot. 1995. "Public Schooling in Rural Pakistan: Efficiently Targeting Girls and a Lagging Region." In D. Van de Walle and K. Nead, eds., *Public Spending and the Poor*. Baltimore, Md.: Johns Hopkins University Press.

ADB (Asian Development Bank). 1997. *Emerging Asia*. Manila.

Baliga, S., S. Goyal, and S. Klasen. 1999. "Education and Marriage Age: Theory and Evidence." Discussion Paper. Department of Economics, University of Munich, Germany.

Barro, R. 1991. "Economic Growth in a Cross-Section of Countries." *Quarterly Journal of Economics* 106(2):407–43.

Barro, R., and J. Lee. 1994. "Sources of Economic Growth." *Carnegie-Rochester Series on Public Policy* 40:1–46.

———. 1996. "International Measures of Educational Achievement." *American Economic Review* 86:218–23.

Barro, R., and X. Sala-i-Martin. 1995. *Economic Growth*. New York: McGraw-Hill.

Berik, G. 2000. "Mature Export-Led Growth and Gender Wage Inequality in Taiwan." *Feminist Economics* 6(3):1–26.

Bloom, D., and J. Williamson. 1998. "Demographic Transition and Economic Miracles in Emerging Asia." *World Bank Economic Review* 12(3):419–55.

Caselli, F., G. Esquivel, and F. Lefort. 1996. "Reopening the Convergence Debate: A New Look at Cross-Country Growth Regressions." *Journal of Economic Growth* 1(3):363–89.

Collier, P., and J. Gunning. 1999. "Explaining African Economic Performance." *Journal of Economic Literature* 37(1):64–111.

Dollar, D., and R. Gatti. 1999. "Gender Inequality, Income, and Growth: Are Good Times Good for Women?" Policy Research Report, Engendering Development, Working Paper No. 1. Washington, D.C.

Dollar, D., and A. Kraay. 2000. "Growth Is Good for the Poor." World Bank, Research Department Washington, D.C.

Drèze, Jean, and Amartya Sen. 1989. *Hunger and Public Action*. New York: Oxford University Press.

Easterly, B., and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4):1203–50.

Forbes, K. J. 2000. "A Reassessment of the the Relationship between Inequality and Growth." *American Economic Review* 90(4):869–87.

Galor, O., and D. N. Weil. 1996. "The Gender Gap, Fertility, and Growth." *American Economic Review* 86(2):374–87.

Heston, A., and R. Summers. 1991. "The Penn World Table Mark 5." *Quarterly Journal of Economics* 106(2):328–68.

Horton, S. 1999. "Marginalization Revisited: Women's Market Work and Pay, and Economic Development." *World Development* 27(3):571–82.

Hill, A. and E. King. 1995. *Women's Education in Development Countries*. Baltimore, Md.: Johns Hopkins University Press.

Klasen, S. 1993. "Human Development and Women's Lives in a Restructured Eastern Bloc." In A. Schipke and A. Taylor, eds., *The Economics of Transition: Theory and Practise in the New Market Economies*. New York: Springer.

———. 1999. "Does Gender Inequality Reduce Growth and Development? Evidence from Cross-Country Regressions." Policy Research Report, Engendering Development, Working Paper No. 7. World Bank, Washington, D.C.

Klasen, S., and Claudia Wink. 2002. "A Turning Point in Gender Bias in Mortality?" *Population and Development Review* 28(2):285–312.

Knowles, S., P. K. Lorgelly, and P. D. Owen. 2002. "Are Educational Gender Gaps a Brake on Economic Development? Some Cross-Country Empirical Evidence." *Oxford Economic Papers* 54(1):118–49.

Lagerlöf, N. 1999. "Gender Inequality, Fertility, and Growth." University of Sydney, Department of Economics, Australia.

Lorgelly, P. K., and P. D. Owen. 1999. "The Effect of Female and Male Schooling on Economic Growth in the Barro-Lee Model." *Empirical Economics* 24(3):537–57.

Lucas, R. 1988. "On the Mechanics of Development Planning." *Journal of Monetary Economics* 22(1):3–42.

Mankiw, N. G., D. Roemer, and D. N. Weil. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107(2):407–37.

Murthi, M., A.-C. Guio, and J. Drèze. 1995. "Mortality, Fertility, and Gender Bias in India: A District-Level Analysis." *Population and Development Review* 21(4):745–82.

NBER (National Bureau for Economic Research). 1998. "The Penn World Tables Mark 5.6." Cambridge, Mass. Available online at www.nber.org.

Palmer, I. 1991. *Gender and Population in the Adjustment of African Economies*. Geneva: International Labour Office.

Pritchett, L., and L. Summers. 1996. "Wealthier Is Healthier." *Journal of Human Resources* 31(4):841–68.

Ravallion, M. 2001. "Growth, Inequality, and Poverty: Looking Beyond Averages." *World Development* 29(11):1803–16.

Roemer, P. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94(5):1002–37.

Sachs, J., and A. M. Warner. 1995. "Natural Resources and Economic Growth." Development Discussion Paper 517a. Harvard Institute for International Development, Cambridge, Mass.

———. 1997. "Sources of Slow Growth in African Economies." *Journal of African Economies* 6(3):335–76.

Schultz, T. P. 1994. "Human Capital, Family Planning, and Their Effects on Population Growth." *American Economic Review* 84(2):255–60.

Seguino, S. 2000a. "The Effects of Structural Change and Economic Liberalization on Gender Wage Differentials in South Korea and Taiwan." *Cambridge Journal of Economics* 24(4):437–59.

———. 2000b. "Gender Inequality and Economic Growth: A Cross-Country Analysis." *World Development* 28(7):1211–30.

Sen, Amartya. 1989. "Women's Survival as a Development Problem." *Bulletin of the American Academy of Arts and Sciences* 43(2):14–29.

———. 1999. *Development as Freedom.* New York: Knopf.

Solow, R. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70(1):65–94.

Standing, G. 1999. "Global Feminization through Flexible Labor: A Theme Revisited." *World Development* 27(3):583–602.

Summers, L. 1994. *Investing in All the People.* Washington, D.C.: World Bank.

Taylor, Alan. 1998. "On the Costs of Inward-Looking Development." *Journal of Economic History* 58(1):1–28.

Thomas, D. 1990. "Intrahousehold Resource Allocation: An Inferential Approach." *Journal of Human Resources* 25(4):634–64.

———. 1997. "Income, Expenditures, and Health Outcomes: Evidence on Intrahousehold Resource Allocation." In L. Haddad, J. Hoddinott, and H. Alderman, eds., *Intrahousehold Resource Allocation in Developing Countries: Models, Methods, and Policy.* Baltimore, Md.: Johns Hopkins University Press.

Tzannatos, Z. 1999. "Women and Labor Market Changes in the Global Economy." *World Development* 27(3):551–70.

UNDP (United Nations Development Programme). Various years. *Human Development Report.* New York: Oxford University Press.

UNICEF (United Nations Children's Fund). 1992. *The State of the World's Children.* New York.

———. 1996. *Women's Indicators and Statistics (WISTAT).* Version 3.0. New York.

Waring, M. 1988. *If Women Counted.* New York: Harper and Row.

World Bank. 1993. *World Development Report 1993: Investing in Health.* New York: Oxford University Press.

———. 2000. *World Development Report 2000/01: Attacking Poverty.* New York: Oxford University Press.

———. 2001. *Engendering Development.* Washington, D.C.

Young, Alwyn. 1995. "The Tyranny of Numbers: Accounting for the Economic Miracle in East Asia." *Quarterly Journal of Economics* 110(3):641–80.

THE WORLD BANK ECONOMIC REVIEW, VOL. 16, NO. 3 375–396

Gender, Time Use, and Change: The Impact of the Cut Flower Industry in Ecuador

Constance Newman

This article uses survey data from Ecuador to examine the effects of women's employment on the allocation of paid and unpaid labor within the household. I compare a region with high demand for female labor with a similar region in which demand for female labor is low. The comparison suggests that market labor opportunities for women have no effect on women's total time in labor but increase men's time in unpaid labor. The increase in men's time in unpaid work reflects women's increased bargaining power in the home.

Economic reforms in Latin America have led to a boom in the growth of nontraditional agriculture exports and along with it a large increase in the demand for rural labor, especially female labor (see Barham and others 1992, Quiroz and Chumacero 1996, Thrupp 1995). The expansion in off-farm employment has provided a needed income source for rural families, and it has had a profound impact on the economic and social fabric of rural communities. Agricultural industries and nontraditional products have grown in direct response to trade and macroeconomic reforms recommended by the World Bank and other institutions. Success stories in Latin America are frequently cited as examples for other countries to follow.[1] The Ecuadorian cut flower industry is a classic example of a growing agricultural export industry in Latin America and one that has a large demand for female labor.

This article investigates two questions about how household time allocation has changed as a result of the flower industry. The first is whether women who

Constance Newman is with the Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture, Washington, D.C Her e-mail address is cnewman@ers.usda.gov The author thanks Harold Alderman, Benedicte de la Briere, Lynn Brown, Jorge Eguiguren, Dean Jolliffe, Nader Kabbani, Shahidur Khandker, Peter Lanjouw, Pilar Larreamendy, Ana Maria Maldonado, Norma Mena, Vicente Merino, Marta Ordoñez, Gioconda Paez, Agnes Quisumbing, Martin Ravallion, Alberto Valle, Parke Wilde, Patricia Zambrano, and especially François Bourguignon and three anonymous reviewers for excellent critiques and suggestions. Funding for this research came from the Gender and Development Thematic Group, the Gender and Poverty Thematic Group, and the World Bank Research Committee.

1 World Bank Country Assistance Strategies typically promote nontraditional agriculture for its high-growth potential See, for example, the Uganda Country Assistance Strategies, *The Challenge of Growth and Poverty Reduction* (World Bank 1996)

DOI. 10.1093/wber/lhf006

work in the flower industry are working harder as a result of combining market labor with unpaid labor at home. The second is whether there is some shifting of responsibility for unpaid labor traditionally performed by women to male household members. Such a change in participation in housework could imply a shift in gender roles or a tradeoff based on relative wages. Several recent studies have examined changes in time use in more industrial countries, and a growing body of literature is looking at similar questions about time use in developing countries.[2] In the United Kingdom, Jenkins and O'Leary (1997) found significant change in the composition of work for men and women, which they attributed to the increased participation of married women in the labor force. Manchester and Stapleton (1991) report similar findings for the United States.[3]

Most work on time use in developing countries has focused on the separate impacts for men and women, with little emphasis on the division of labor in the home and how it may change over time. An exception is a recent article by Fafchamps and Quisumbing (1998), which looks at the static determinants of the intrahousehold division of labor in Pakistan. They found that gender, family status, and human capital variables were all significant determinants, but they did not address the question of how or why the division of labor may change over time.

People often assume that gender roles are fixed in developing economies. Although it is true that roles in many developing societies are more narrowly defined for women, pressures from modernization are provoking swift changes. Married women's participation in paid labor has risen rapidly around the world, especially in export niches.

In Ecuador the flower industry is only 10 years old. Before it developed, women in these rural areas had little (if any) paid employment. Employment in flowers has been one of the first types of paid off-farm employment offered to women and the only employment offered to women in large numbers.[4] This kind of change in women's market participation is likely to induce some changes in household time allocation.

The challenge is to measure the dynamic effect of household labor supply decisions on the reallocation of responsibilities within the home when time-series data are not available and when, from a static modeling perspective, time allocation decisions must be assumed to be made simultaneously. One solution

2 For developing economies, see Alderman and Chishti (1990), Fafchamps and Quisumbing (1998), Ilahi (1999), Khandker (1988), and Skoufias (1993).

3. See also Bittman (1999) for a comparison of trends in Finland and Australia and van der Lippe and Siegers (1994) for the Netherlands. Bittman cautions that although gender differentials in housework in more developed economies are declining, men devote half the time that women do to unpaid work.

4. Firm managers and many others say that women are more efficient than men in the detail-oriented work required in the flower industry. Others assert that women are hired because they are willing to work for lower pay than men are. It is most likely some combination of factors: productivity, labor supply, and the historical association of women with postharvest agricultural work.

would be to model male and female decisions separately and treat women's income (or market participation) as exogenous to men's decisions to carry out housework. This is the working assumption in many noncooperative models, and it was a common assumption in early models of labor supply. For example, it was common to include husband's income as an exogenous determinant of female labor supply, because it was assumed—probably correctly for some time—that a husband's participation was independent of a wife's participation. For the purposes of this article, a separable model would negate the hypothesized relation between women's market work and male housework, an effect that is believed to take place over time.

The framework for this analysis is a model in which time use decisions in all activities are simultaneous and interlinked across household members. Directly estimating such a system of decisions is complicated by the large differences across households found in most developing societies, both in the number of people per household and in the types of relatives who live together. The effect of women's work on housework allocations is captured indirectly in two ways. First, time allocation outcomes are compared in two different "states of nature," one in a region in which flowers are produced and the other in a region where they are not. Second, reduced-form determinants of time use are estimated together with wages and a dummy variable for the flower-producing region.

I. The Conceptual Framework

Building on the framework proposed by Browning and Chiappori (1998), I characterize household welfare as a weighted sum of individual utilities. This approach follows in the tradition of modeling the household decision as a collective decision rather than assuming the household has one representative, or "unitary," utility function.

Let each member's unique utility contribute to household choices depending on their relative weight in the household, written as

$$U = \sum_{i=1}^{n} \mu^i U^i, \qquad \text{where} \quad \sum_{i=1}^{n} \mu^i = 1 \tag{1}$$

The choice functions that result from maximizing equation (1) subject to a budget constraint depend on the weights μ^i. However, as in the demand example described by Browning and Chiappori, we do not observe choice responses to changes in the μ's. We observe only choices made at one set of μ values for any given price-income bundle. Therefore, the weights can be written as functions of the parameters that determine the choices in question. Browning and Chiappori call these parameters "preference factors" and distinguish them from other factors that may also be included, such as changes in divorce laws or regional policy differences, which they call "distribution factors."

Expanding the basic model to show μ as a function of preference and distribution factors, we designate household welfare and the budget constraint by

$$U = \sum_{i=1}^{n} \mu^i (w^1, \ldots, w^n, W; H^1, \ldots, H^j) U^i (C^i, R^i) \tag{2}$$

$$\sum_{i=1}^{n} C^k - \sum_{i=1}^{n} w^i L^i = W \tag{3}$$

where U^i is the utility of individual i defined over the numeraire good (C^i) and leisure (R^i) consumed by individual i. Leisure comprises rest and recreation and is expressed by $R^i = T^i - L^i - D^i$, where T^i is the total time endowment, L^i represents paid labor, and D^i represents domestic labor.[5] The individual utility weights, $\mu^i(\cdot)$, sum to one over the n members of the household and differ across households by the household wealth, W, and the j household characteristics in the vector H. The utility weights differ across household members by the wage they earn in market labor, w^1. Domestic labor is a necessary service that has to be conducted by at least one member of the family and is defined by the constraint

$$\sum_{i} D^i = D_m \tag{4}$$

where D_m is the amount of time that needs to be allocated to domestic work in the mth household. Maximizing household welfare (equation [2]) subject to equations (3) and (4) and imposing nonnegativity constraints on individual labor supply and domestic work supply, $L^i \geq 0$, and $D^i \geq 0$, results in a series of labor and domestic work supply functions:

$$L_a^i = f(w^1, \ldots, w^n, W, \mu[w^1, \ldots, w^n, W; H^1, \ldots, H^j], \lambda^k) \tag{5}$$

for $i = 1, \ldots, n$ members of the M households, all j household characteristics; $k = 1, \ldots, 4$ Lagrange and Kuhn-Tucker multipliers, and where a is for paid labor and domestic work supply functions.

As equation (5) shows explicitly, wages affect the labor and domestic work supply functions directly as well as indirectly through the distribution function μ. The effect on μ can be thought of as a bargaining effect. As wage opportunities change among household members, the supply outcomes of each one can be affected beyond the traditional substitution and income effects of a wage change. Suppose one family member's wage increases relative to the others. The impact on the distribution function is unclear and would depend on the rules that determined the function. It is possible that the higher the wage earned by an individual, the higher his or her utility would be weighted, because the higher wage

5. For simplicity of presentation, home production is not included, because the focus is on the underlying structure of paid labor versus domestic labor supply decisions. Home production could be added, but it would lead to complications, as pointed out by Apps and Rees (1997) and discussed by Chiappori (1997). To avoid the pitfall associated with leaving it out that Chiappori notes—namely, that work in the home may accidentally be treated as leisure—all hours of work, whether in the home or elsewhere, are counted as either paid or domestic (unpaid) labor. This is possible because detailed information on time use by activity is available.

increases the person's ability to survive independently away from the household. However, a higher wage would have an ambiguous bargaining effect on labor supply, depending on the relative value that the individual places on consumption versus leisure.

The bargaining effect of a higher relative wage would have a clear impact on domestic work supply ("housework"). Because housework is not paid, the person would choose more market work or more leisure but never more housework. An exception would be the case in which the work itself produces positive utility for the individual. That may be true for housework that involves taking care of other family members, but then most forms of work have some aspects that provide positive utility. Overall, a higher wage would translate into less housework for most individuals through a bargaining effect.

II. The Data

The survey was conducted in two regions of northern Ecuador, both around the small towns of Cotocachi and Cayambe. The two regions were chosen for their cultural and ecological similarities. Each region contains a relatively concentrated regional center in the lowlands and rural hamlets in the neighboring hillsides. The regions are about 200 km apart; Cotocachi is two hours and Cayambe is three hours by car from the capital city, Quito.

The data were collected in May and June 1999 as part of a larger study of the effects of flower development on household resource allocation decisions. The survey was designed specifically to answer the questions raised in this article on time use as well as to answer questions about intrahousehold consumption and production decisions. In total 558 households were surveyed, resulting in 2,541 individual observations from all members of each family (1,861 individuals were 10 years or older). The survey is modeled after the World Bank's Living Standards Measurement Survey and uses many of the same modules used in living standards measurement surveys conducted by the Ecuadorian government. The same team of Ecuadorian consultants that carried out the national surveys conducted the survey for this project. Two Ecuadorian anthropologists conducted a complementary series of participatory workshops, focus groups, and individual interviews in the same two regions using the same sample design.

The survey data include detailed modules on expenditures, economic activity (including agriculture and small businesses), health, education, fertility, credit and savings, and time use. Two types of time use data were collected because of their different strengths.[6] The 24-hour data are considered by many to be more accurate because they are more detailed and it is easier for a respondent to remember what was done the day before. But 24-hour data are more likely to miss

6 See Juster and Stafford (1991) and Robinson and Gershuny (1994) for a full discussion of time use measurement issues.

unusual or irregular activities. Because men's contribution to housework may be an example of an irregular activity or one that is not done daily, we also asked for time dedicated to housework, rest, recreation, and work each day the previous week. Weekly data of this nature have the disadvantage of being less precise and more subject to recall error, but they have the advantage of being less burdensome to the interviewee. The 24-hour recall data were collected only for the male and female heads of household. The weekly data were collected for all household members interviewed. In the flower industry region the sample was also stratified by whether or not the wife or female household head works in the flower industry.

A classic problem arises when comparing two sites, or two experimental groups, when the effect being measured is more (or less) likely to occur in the treated group for reasons unrelated to the treatment. This problem would arise here if the location choice of the flower industry entrepreneurs were correlated with qualities of the workforce that also influence workers' time allocation decisions. According to flower producers, the characteristics of the workforce are irrelevant to their location choice, which was guided by the unique combination of microclimatic characteristics in Cayambe. They claim that anyone can be trained in a short period to do the work required. One may argue that in reality the flower producers choose a more productive workforce—younger and more educated—and that more productive and less productive workers allocate their time between market and domestic work differently. But most of these kinds of worker characteristics can be controlled for in the analysis.

The main impact of the flower industry in the Cayambe region has been to offer much more and better paid employment for women. Median wages paid to women in the industry are twice those paid in other industries, even in other industries in the Cayambe region.

Before the flower industry moved into the area in the late 1980s, Cayambe had been a center of dairy production and milk-related products. The "control" area, Cotocachi, resembles Cayambe as it was about 15 years ago. It has a stable economy with a small artisan industry in leather products. As a result of the stability of the local economy in Cotocachi, few people migrate to the flower-producing region, and women are much less likely than men to migrate at all from Cotocachi.

Why have the two labor markets not equilibrated? Why would the residents of Cotocachi not move or even commute to work in the flower industry given the higher wages offered there? For men the jobs in the flower industry are not clearly better paying than the jobs in Cotocachi. For women the main reason they do not migrate is the cultural importance of staying close to home, coupled with the fact that their families are not excessively poor. In the qualitative work that accompanied the survey, we found that Cotocachi women did not view flower employment as an option. They said either that their husbands would not allow them to work or that the work would be detrimental to family relations. Here is a typical explanation from a Cotocachi woman:

> He didn't want me to go out [to work]. . . He refused to let me. . . . It was a problem when I worked [before]. . . His family wouldn't talk to me, [they would say] "Why go outside to work when you can work in the house itself?" They talked about it so much even my own brothers and sisters were asking, "Why are you working?"

Women from Cayambe faced similar pressures when they first wanted to work in the flower industry. Almost all of the married women interviewed said their husbands were initially opposed to them working. But once they started earning money, opposition decreased and their economic contributions were appreciated. Many of the younger women said their parents had been opposed to the idea. Many parents wanted their daughters to stay in school. Here are quotations from Cayambe women about their decisions.

> He told me "no." . . . but the decision was mine.

> In my case, my parents did not want me to work, but [because of] the economy, you had to because there wasn't enough money to pay for school.

For women in Cayambe who did not work and who participated in the quantitative survey, more than half said they were not working because of family opposition.[7]

The main difference between the inhabitants of Cotocachi and the migrants to Cayambe is in the wage opportunities for men. The migrants went in search of employment for both the men and the women of the family. Cotocachi men earned wages similar to those offered by the flower industry (more on wages later), and it was not socially acceptable for the women to travel alone. For these reasons the two markets for women's labor did not equilibrate.

Cayambe and Cotocachi have different microclimatic characteristics but similar ecological features, cultures, and histories of land tenure (table 1). Descriptive statistics give only a limited comparison of the two areas, because we would expect to see differences in Cayambe as a result of the flower industry. The main differences are the age profiles and the proportion of migrants to the area. The population of Cayambe is generally younger and has a higher proportion of migrants. Most of the other basic characteristics, including education levels, marriage patterns, household composition, religious affiliations, and other organizational affiliations, are similar.

The large proportion of migrants in Cayambe raises the question of how the migrants affect the comparability of the groups. Migrants to Cayambe made up 26 percent of the sample population in Cayambe and only 6 percent of the sample in Cotocachi. The migrants in Cayambe were from all over the country, but a large number were from the poorer areas along the coast. Migrants were more likely to be young adults. With fewer elderly in the migrant sample, a higher

7. See Newman and others (forthcoming) for general findings from the quantitative and qualitative studies.

TABLE 1. Demographics of Cayambe and Cotocachi Areas

Item	Cayambe all		Cayambe migrants		Cotocachi	
	Number	Percentage	Number	Percentage	Number	Percentage
Sample size	1,916	75	497	26	625	25
Age group						
13 and younger	684	35.7	66	13.3	209	33.4
14–19	260	13.6	88	17.7	74	11.8
20–25	298	15.6	139	28.0	55	8.8
26–35	309	16.1	129	26.0	73	11.7
36–45	147	7.7	49	9.9	60	9.6
45–55	97	5.1	14	2.8	51	8.2
56–65	67	3.5	8	1.6	43	6.9
66 and older	54	2.8	4	0.8	60	9.6
Education						
None	137	8.7	10	2.0	59	10.6
Nursery	10	0.6	2	0.4	0	0
Preprimary	52	3.3	1	0.0	14	2.6
Basic education	5	0.3	0	0.2	0	0
Primary	834	52.9	256	51.5	300	54.8
Secondary	460	29.2	204	41.1	138	25.2
Superior university	72	4.6	21	4.2	33	6.0
Superior not university	6	0.4	3	0.6	2	0.4
Graduate	0	0.0	0	0.0	1	0.2
Marital status						
Free union	163	13.7	109	26.1	27	6.7
Married	501	42.0	147	35.3	198	49.0
Single	435	36.5	138	33.1	132	32.7
Separated	41	3.4	15	3.6	10	2.5
Divorced	13	1.1	4	1.0	7	1.7
Widowed	40	3.4	4	1.0	30	7.4
Female head of household	95	23.1	33	19.5	37	25.3
Household with children younger than 15 years old	684	35.7	66	13.3	209	33.4
Households with children younger than 6 years old	340	17.7	0	0.0	78	12.5
Relation to head of household						
Household head	412	21.5	169	34.0	146	23.4
Spouse	284	14.8	108	21.7	99	15.8
Son or daughter	892	46.6	121	24.4	297	47.5
Son- or daughter-in-law	33	1.7	10	2.0	11	1.8
Niece or nephew	157	8.2	6	1.2	44	7.0
Mother or father	13	0.7	4	0.8	5	0.8
Mother- or father-in-law	6	0.3	2	0.4	6	1.0
Brother or sister	51	2.7	38	7.7	7	1.1
Brother- or sister-in-law	15	0.8	11	2.2	2	0.3
Other relatives	36	1.9	18	3.6	5	0.8
Other	17	0.9	10	2.0	3	0.5

Note: A large-sample rank sum test was performed on the Cayambe and Cotocachi samples, and the only variable distributions in the table for which the hypothesis that they are the same was rejected were "age group" and "relation to head of household." The differences in "relation to head of household" across Cayambe and Cotocachi, however, are quite small, reflecting the strength of the test because of its large sample properties.

Source: Surveys conducted in May and June 1999; see text.

percentage of the sample had secondary education. Migrants had smaller families, fewer children, and fewer female-headed households. Perhaps the migrants had cultural norms that made them more open to gender role changes. Because of these differences, migrant status is separately included in the econometric estimations.

III. Time Use Outcomes: How Do Allocations Differ in Cayambe and Cotocachi?

Time use in the major daily activities for the past 24 hours is shown in table 2. The total time worked by women in Cayambe was slightly less than that worked by women in Cotocachi, although the differences are not significant. Compared with men, women in both areas spent significantly more time working, including both paid work and housework.[8] The ratio of men's time in total work to women's was only slightly higher in Cayambe (82 percent) than in Cotocachi (80 percent). Men worked about 8.5 hours a day, and women worked about 10.5 hours a day—a difference that is common in developing economies (Ilahi 1999, World Bank 2001). Men's and women's total working hours are much more similar in industrial countries.

Not surprisingly, women in Cayambe spent more time performing paid work (229 minutes, or 3.8 hours, a day) than women in Cotocachi (171 minutes, or 2.9 hours, a day), but women in both regions spent less time performing paid work than men. Given the relatively short history of women's participation in paid work, it is not surprising that women worked less than men, but this is not representative of the population, because the survey deliberately included women who did not work in the sample design. Men in Cayambe spent significantly more time performing paid work (361 minutes, or 6 hours, a day) than men in Cotocachi (302 minutes, or 5 hours, a day).

Women did most of the housework in both Cayambe and Cotocachi (table 2). Cayambe women spent an average 327 minutes (5.5 hours) a day on housework, whereas men averaged 62 minutes a day of housework. In Cotocachi women spent 353 minutes (5.9 hours) on housework, and men spent 57 minutes. Men did slightly more housework in Cayambe than in Cotocachi, and women did slightly less.

Much larger differences emerge when looking at marital status and labor market participation (table 3). Married male household heads who worked in Cotocachi performed an average of 31 minutes a day of housework, whereas married male household heads who worked (in any job) in Cayambe performed an average of 57 minutes of housework a day. Both sets of married men did more

8. All reported differences are statistically significant unless reported otherwise. However, in subgroups such as married people, strictly speaking the sample was not designed to make these statistical comparisons. There are too few observations to correct the standard errors in several cases of the more detailed subgroup comparisons.

TABLE 2. Use of Time, by Gender and Marital Status
(minutes per day)

Activity	Men	Women	Married men	Single men	Married women	Single women
Cayambe						
Farm work	60.26[a]	51.42[a]	62.95[a]	31.04	52.51[a]	48.91
Paid work	360.69[a,b]	228.67[a,b]	364.64[a]	317.71	226.94[a,b]	232.62
Community work	13.09	4.42	13.26	11.25	5.30	2.42
Housework	61.79[b]	327.11[b]	62.90[a,b]	49.67[a,b]	331.14[b]	317.90[b]
Total work	516.48[b]	626.90[b]	524.40[b]	430.20[b]	631.33[b]	616.81[b]
Recreation	100.41[b]	66.44[b]	96.13[b]	146.88[b]	70.36[b]	57.50[b]
Personal care	328.06[a,b]	249.41[a,b]	325.10[b]	360.21[a,b]	243.90[a,b]	261.98[b]
Total time	944.94[a]	942.76[a]	945.64[a]	937.38	945.59[a]	936.30
Number of observations	298	411	270	28	286	125
Cotocachi						
Farm work	116.01[a]	79.31[a]	118.14[a]	96.00	81.47[a]	74.05
Paid work	302.45[a,b]	171.02[a,b]	311.65[a,b]	216.00	166.88[a,b]	181.07
Community work	13.85	11.46	15.32	0.00	14.71	3.57
Housework	57.36[b]	353.45[b]	38.94[a,b]	230.50[a]	369.87[b]	313.57
Total work	507.26[b]	630.38[b]	499.84[b]	577.00	648.87[b]	585.47
Recreation	108.17	74.10	103.72	150.00	68.68	87.26
Personal care	289.52[a,b]	204.56[a]	299.95[b]	191.50[a]	191.82[a,b]	235.48
Total time	904.95[a]	909.03[a]	903.51[a]	918.50	909.36[a]	908.21
Number of observations	110	144	99	11	102	42

	Cayambe			Cotocachi		
Men/women ratio	All	Married	Single	All	Married	Single
Farm work	1.17	1.20	0.63	1.46	1.45	1.30
Paid work	1.58	1.61	1.37	1.77	1.87	1.19
Community work	2.96	2.50	4.65	1.21	1.04	0.00
Housework	0.19	0.19	0.16	0.16	0.11	0.74
Total work	0.82	0.83	0.70	0.80	0.77	0.99
Recreation	1.51	1.37	2.55	1.46	1.51	1.72
Personal care	1.32	1.33	1.37	1.42	1.56	0.81

Note: Figures are based on 24-hour data. Total work includes time spent performing other activities (fetching water, selling and repairing homes).

[a]Significantly different from that of the opposite area at 95 percent confidence.

[b]Significantly different from that of the opposite gender at 95 percent confidence.

Source: Surveys conducted in May and June 1999; see text.

housework when their wives worked, but men in Cayambe in both categories spent significantly more time doing housework than men in Cotocachi.

Table 3 also shows a large flower sector impact. Married male household heads who worked in flowers did more housework than those who worked in other sectors (69 minutes versus 47 minutes), and this difference grew when their wives also worked in the sector. Married and working male household heads did the most housework of any group of men (71 minutes) when they worked in flowers and their wives worked (mostly in flowers). Married male household heads

TABLE 3. Time Spent Performing Household Tasks, by Gender, Marital Status, and Labor Market Participation

Item	Cayambe			Cotocachi			Difference in means *t*-test[a]
	Number	Mean	SE	Number	Mean	SE	
Male household head, married[b,c]							
Works	256	57.41	5.36	93	31.18	8.07	0.0050
And wife works	168	62.49	6.56	44	31.93	11.75	0.0160
And wife doesn't work	88	47.73	9.24	49	30.51	11.22	0.1258
Works in flowers	120	68.79	8.02*[c]	93	31.18	8.07	0.0007
And wife works	99	70.76	8.66	44	31.93	11 75	0.0058
And wife doesn't work	21	59.52	21.18	49	30.51	11.22	0.0960
Doesn't work in flowers but works	136	47.38	7.10*	93	31.18	8.07	0.0690
And wife works	69	50.62	9.93	44	31.93	11.75	0.1164
And wife doesn't work	67	44.03	10.21	49	30.51	11.22	0.1897
Female household head, single[c]							
Works	86	276.48	22.59	33	274.70	27.11	0.5176
Works in flowers	51	199 84	23.10*	33	274.70	27.11	0.0205
Doesn't work in flowers	35	388.14	37.05	33	274.70	27.11	0.9914
Wife of household head[b,c]							
Works	242	291.09	12.69	86	354.91	19.58	0.0045
And husband works	232	287.88	12.99	83	359.36	19.88	0.0021
Doesn't work	44	528.86	33.47	16	450.31	42.12	0.8980
Works in flowers	145	232.49	12.45*	86	354.97	19.58	0.0000
Works in flowers and husband works	143	230.92	12.53*	83	359.36	19.88	0.0000
Doesn't work in flowers, but works	97	378.69	22.96*	86	354.91	19.58	0.7812
And husband works	89	379.42	24.37*	83	359.36	19.88	0.7360

*Significant difference at 95 percent or more between those in the treatment group working in flowers and those working in other sectors.

[a]The tests for different means are $\mu^a > \mu^b$ for men and $\mu^a < \mu^b$ for women

[b]The total number of married male household heads does not exactly match the number of wives of male household heads due to missing responses (there are 271 wives and 267 husbands).

[c]Not all subcategories are shown. The subcategories that are missing have too few observations to perform tests on their means.

Source: Author's calculations based on surveys conducted in May and June 1999; see text.

who worked in another sector and whose wives worked spent 51 minutes a day on housework. Although not significant, men who worked but whose wives did not spent more time on housework in Cayambe than in Cotocachi. This is especially interesting because the higher values for Cayambe men when women do not work is clearly not a substitution or income effect. Overall, these data suggest that men in Cayambe do more housework whether their wives work or not, but especially so if their wives work and if the men themselves work in the flower industry.

What is it about the flower sector that would induce higher levels of unpaid work by men? One possible explanation is wage differences. Men were paid more

than women on average (and at the median) in both Cayambe and Cotocachi, and this was also true of the subgroup of flower industry workers in Cayambe (table 4). However, the differences between men and women were much smaller among workers in the flower industry, with men earning 5,626 sucres and women averaging 5,552 sucres. These figures were much closer than the average figures in other sectors in Cayambe, where men average 5,817 and women average 3,704 sucres. Among married couples who worked in the flower industry, women actually earned more than men did.

Among working women, married women in Cotocachi did much more housework (355 minutes, or 5.9 hours, a day) than married women in Cayambe (291 minutes, or 4.9 hours, a day) (see table 3). The impact of the flower sector was quite strong. Among married Cayambe women who worked, those who worked in flowers spent much less time doing housework (232 minutes, or 3.9 hours a day) than women in other sectors (379 minutes, or 6.3 hours). In fact, the average time spent on housework by Cayambe women working in other sectors was more than that of women in Cotocachi. The flower impact is also seen in the amount of housework done by wives employed in the flower sector whose husbands also worked in the sector (222 minutes, or 3.7 hours, a day) compared with women whose husbands worked in other sectors (253 minutes, or 4.2 hours, a day). These statistics are symmetrical to the findings for men and can also be explained by the relatively higher wages for women in the flower industry.

The weekly data (see table 5) reveal the same patterns as the 24-hour data. What is evident in the weekly data, which are more representative for single people be-

TABLE 4. Wages by Gender, Marital Status, and Work Type

	Men				Women			
	N	Mean	Median	SD	N	Mean	Median	SD
Cotocachi								
All	141	5,668	3,000	8898	118	3,451	2,112	4,585
Married	101	6,454	3,000	10175	53	3,340	2,308	2,697
Cayambe								
All	444	5,715	4,568	5912	479	4,977	4385	5,719
Married	297	5,966	4,686	6263	225	5,687	4,761	6,656
All who work in flowers	236	5,626	4,823	3880	330	5,552	4,630	5,833
Married who work in flowers	157	5,913	4,938	4290	163	6,263	4,941	6,756
All who work in other sectors	208	5,817	3,846	7595	149	3,704	1,978	5,257
Married who work in other sectors	140	6,025	4,075	7929	62	4,173	2,464	6,182

Source: Author's calculations based on surveys conducted in May and June 1999; see text.

TABLE 5. Average Hours per Week Spent Performing Main Activities by Gender and Marital Status

Activity	Men	Women	Married men	Single men	Married women	Single women
Cayambe						
Paid work	39.1[a]	33.5[a,b]	50.6[a]	36.3[a]	25.9[a,b]	31.3[a,c]
Housework	7.0[a]	24.0[a]	7.8[a,c]	30.8[a,c]	6.1[a]	18.7[a]
Recreation	15.3[a,b]	12.8[a]	12.8[c]	11.5	18 2[a]	13.9[a]
Sleep	56.3[b]	56.6[b]	55.1[b]	55.9[b,c]	57.7	57.1[b]
Cotocachi						
Paid work	36.9[a]	28.5[a,b]	53.5[a]	31.7[a]	18.5[a,b]	26.1[a,c]
Housework	6.5[a]	24.4[a]	5.9[a,c]	34.2[a,c]	7.0[a]	17.2[a]
Recreation	13.6[b]	12.7	11.2[c]	11.6	16.3[a]	13.5[a]
Sleep	58.3[b]	58.5[b]	57.3[b]	57.8[b,c]	59.5	59.0[b]

	Cayambe			Cotocachi		
Men/women ratio	All	Married	Single	All	Married	Single
Paid work	1.17	1.39	0.83	1.29	1.69	0.71
Housework	0.29	0.25	0.33	0.27	0.17	0.41
Recreation	1.20	1.11	1.31	1.07	0.97	1.21
Sleep	0.99	0.99	1.01	1.00	0.99	1.01

Note: Figures are based on weekly data. Sum includes time in other activities (fetching water, selling and repairing homes).

[a]Significantly different from that of the opposite gender at 95 percent confidence.

[b]Significantly different from that of the opposite area (Cayambe or Cotocachi) at 95 percent confidence.

[c]Significantly different from that of the opposite area (Cayambe or Cotocachi) at 90 percent confidence.

Source: Author's calculations based on surveys conducted in May and June 1999; see text.

cause of the way they were collected, is that single men in Cayambe and Cotocachi spend almost the same time performing housework. This makes the difference among married men even more remarkable.

IV. DETERMINANTS OF TIME USE BY GENDER

This section presents an econometric analysis of all the factors that may determine time use. Potential determinants include individual characteristics, such as age, education, and marital status, as well as household and regional characteristics. One goal of the analysis is to differentiate the wage-related substitution effects from the bargaining effects specified in the distribution function $\mu(\cdot)$. Men in Cayambe spent more time performing housework. Was this merely because of wage differences, or was it also because women gained bargaining power over time as a result of their access to regular wage employment? Intuitively, it is clear that there was more than just a pure wage–leisure tradeoff involved, given that women's total time in both domestic and paid work far exceeded that of men. If

economic forces were the only forces at work, we would have to conclude that women preferred much less leisure than men preferred, even in Cayambe, where their market wages were similar. As with the imbalance in leisure, the imbalance in housework responsibilities by gender is a social norm. The most likely way for that norm to change is through women's increased bargaining power gained from market wages.

It is striking to hear how some women in Cayambe described the changes they have experienced:

> When my husband says to me, "Why don't you iron the clothes?" I say, "You too grab the iron and iron." I say, "I don't have time." He keeps quiet, he doesn't say anything to me, he knows he can't change anything.

The bargaining effects of wages on time use are difficult to separate from the substitution effect of wages. To the extent that wages are influencing behavior through a bargaining effect in this analysis, the Cayambe dummy could capture the effect, because the two sample sites are very comparable except for the presence of the flower industry. Own wages and spouses' wages are used together with the Cayambe dummy to help separate the two effects. The wages would be more likely to capture the substitution effects because theoretically they have the most direct impact on utility. The Cayambe dummy is more likely to capture the bargaining effect, because the whole region would be subject to the changing social norms.

Three models were used to test the hypotheses, the most appropriate of which was the censored least absolute deviation (CLAD) model. A Heckman model was originally chosen to correct for self-selection bias. With the Heckman model, one must specify at least one or two variables as determinants of either a person's participation in an activity or the amount of time he or she spent in the activity. There are no theoretical guidelines for making this choice, however, and the results can differ significantly depending on the choice made. The Tobit model, a single-equation model in which the censoring of values at zero is accounted for in the estimated function, also requires the errors to be normally distributed. Tests for normality, looking at skewness and kurtosis, failed.[9]

The CLAD model has the advantage of not requiring any assumptions about the distribution of the errors.[10] Consistent estimates for the β vector are obtained by minimizing

$$\sum_{i=1}^{n} \left| y_i - \max(0, x_i\beta) \right| \tag{6}$$

9. Several models were presented in earlier versions of this article, and the results were all qualitatively similar. An earlier article and other estimates are available on request.

10. See Deaton (1997) for a detailed explanation of this model, proposed by Powell (1984), and for the estimation method proposed by Buchinsky (1994). The method consists of iterative median regressions (in which the observations for which the predicted values are less than zero are discarded until all predicted values are positive) and standard errors are based on bootstrap estimates.

where y_i is measured as the share of paid or unpaid labor and x_i represents a vector of all determinants including the wages and the Cayambe location dummy.

The disadvantage of using this estimator compared with the Tobit is the reduced precision: The estimation method uses a smaller sample, and in a test by Deaton (1997) the standard errors were found to be one and a half times larger than those from Tobit. In larger samples, such as 1,000, according to Deaton, the bias-variance tradeoff becomes more favorable. Since this sample is relatively small, the Tobit results are also presented from one of the main housework models (tables 6 and 7).

The econometric model estimated for the Tobit is written as

$$\begin{aligned} y_i^* &= \beta_0 + \beta_x' x_i + \beta_w w_i + \beta_h' h_i + \varepsilon_i, \\ y_i &= 0 \quad \text{if } y_i^* \leq 0, \\ y_i &= y_i^* \quad \text{if } y_i^* > 0. \end{aligned} \tag{7}$$

In this equation y is measured as the share of unpaid labor, x_i represents a vector of individual characteristics, w_i represents the individual's own hourly wage, and h_i represents a vector of household characteristics for individual i, including the Cayambe location dummy.

The determinants of time use were estimated separately by gender, because the main goal is to understand how male and female behaviors differ not from each other but from members of the same gender. The parameters affecting time allocation for men and women are assumed to differ, particularly for time allocation in housework. Because the households are not homogeneously structured, the share of each individual's contribution to total household labor or domestic work is estimated.

Hourly wages were included in the models because of the importance of testing for a bargaining effect. Predicted wages were used to include nonworking household members. The predicted wage was estimated as a function of age, education, urban location, and variables thought to affect the wage but not labor allocation. These variables included mother's education, tenure on the job, and the ethnic background of the household (assuming there is some level of discrimination against indigenous household members).

Another complication is that the model should include the wages of other relatives in the household or some measure of relative wages. This was found to have serious limitations because of the heterogeneity across households. Different sizes, compositions, and numbers of adults who work make it impossible to have representative wages for each type of family relation. To reduce this problem and others related to heterogeneous households, we estimated wages separately for wives and husbands. The wife's or husband's wages were included in those estimates along with the individual's own wage. Only one household member is included in these models to reduce errors associated with possible correlation across family members. The weekly time use data were used because of the larger sample size.

TABLE 6. CLAD and Tobit Estimates of Men's Share of Time Performing Housework and Paid Work
(dependent variable = individual's share of housework)

Item	Housework			Paid work	
	Tobit All men	CLAD All men	CLAD Married household heads[a]	CLAD All men	CLAD Married household heads[a]
Age/10	−0.0 (−0.09)	−0.03 (−0.94)	−0.07* (−1.82)	0.25* (7.11)	0.20* (3.55)
Age squared/100	0.0 (0.09)	0.0 (0.78)	0.0 (1.52)	−0.0* (−7.14)	−0.0* (−3.88)
Education	−0.002 (−0.59)	0.007* (2.00)	0.004 (0.85)	−0.006 (−1.12)	0.006 (0.92)
Age difference between wife and husband	n.a.	n.a.	0.002 (1.09)	n.a.	0.002 (0.75)
Education difference between wife and husband	n.a.	n.a.	0.000 (0.03)	n.a.	0.014 (1.57)
Married	0.061* (2.10)	0.052* (1.89)	n.a.	0.076* (2.90)	n.a.
Widowed, divorced, or separated	0.216* (3.78)	0.161 (1.14)	n.a.	0.112 (1.01)	n.a.
Own wage (thousands of sucres)	−0.006 (−0.92)	−0.007 (−1.31)	−0.003 (−0.46)	0.032* (4.05)	−0.004 (−0.34)
Wife's wage (thousands of sucres)	n.a.	n.a.	−0.001 (−0.81)	n.a.	−0.000 (−0.00)
Household size	−0.050* (−7.57)	−0.038* (−6.00)	−0.022* (−2.35)	−0.056* (−9.43)	−0.086* (−10.34)
Number of children	0.045* (4.65)	0.029* (3.77)	0.015 (1.31)	0.049* (6.11)	0.091* (6.85)
Ratio of females to males in household	−0.038* (−3.01)	−0.037* (−3.50)	−0.029 (−1.63)	0.005 (0.39)	0.019 (1.04)
Household assets(thousands of sucres)	−0.0002* (−2.19)	−0.0002* (−2.20)	−0.0001 (−1.40)	−0.0002 (−0.79)	0.000 (0.67)
Urban location	0.080* (3.68)	0.060* (2.85)	0.077* (3.32)	−0.066* (−2.47)	−0.022 (−0.78)
Cayambe region	0.120* (5.34)	0.128* (5.28)	0.056* (2.33)	−0.029 (−1.47)	−0.064* (−1.82)
Migrant in Cayambe region	0.015 (0.73)	0.022 (1.43)	0.048* (2.32)	0.022 (1.05)	−0.015 (−0.71)
Number of observations	789	539	305	701	360
Pseudo R^2	0.3344	0.1169	0.1248	0.2895	0.1971

n.a. Not applicable.
*Significant at the 95 percent level.
[a]Only one person per household is selected for this model to control for correlation within the household.
Note: *t*-statistics are shown in parentheses.
Source: Author's calculations based on surveys conducted in May and June 1999; see text.

TABLE 7. CLAD and Tobit Estimates of Women's Share of Time Spent Performing Housework and Paid Work (dependent variable = individual's share of housework)

	Housework			Paid work	
Tobit	CLAD	CLAD	CLAD	CLAD	
Item	All women	All women	Married women[a]	All women	Married women[a]
Age/10	0.22* (8.59)	0.25* (6.91)	0.10 (0.98)	0.71* (8.49)	-0.03 (-0.42)
Age squared/100	-0.0* (-7.72)	-0.0* (-6.13)	-0.0 (-1.09)	-0.0* (-8.99)	-0.0 (-0.39)
Education	0.005 (1.46)	0.007 (1.27)	0.001 (0.11)	0.009 (1.56)	-0.011* (-2.25)
Age difference between wife and husband	n.a.	n.a.	0.004 (1.56)	n.a.	-0.005* (-2.48)
Education difference between wife and husband	n.a.	n.a.	-0.004 (-0.30)	n.a.	-0.026* (-3.08)
Married	0.126* (6.20)	0.153* (7.30)	n.a.	-0.079* (-2.05)	n.a.
Widowed, divorced, or separated	0.097* (3.18)	0.114* (2.38)	n.a.	0.240* (3 20)	n.a.
Own wage (thousands of sucres)	-0.020* (-2.62)	-0.024* (-2.03)	-0.015 (-0 76)	-0 023* (-1.70)	0.011 (1 05)
Husband's wage (thousands of sucres)	n.a	n.a.	0.001 (0.62)	n.a.	-0.001 (-0.32)
Household size	-0.083* (-16.26)	-0.091* (-14.06)	-0.082* (-5.26)	-0.098* (-9.22)	-0.099* (-8.21)
Number of children	0.072* (9.48)	0.070* (9.14)	0.061* (3.06)	0.077* (5.33)	0.082* (4.98)
Ratio of females to males in household	-0.051* (-6.09)	-0.056* (-5.20)	-0.050* (-2.07)	0.034* (2.51)	0.022 (1.51)
Household assets (thousands of sucres)	-0.000 (-0.60)	-0.000 (-1.16)	0.000 (0.19)	-0.000 (-0.58)	0.000 (0.46)
Urban location	-0.046* (-2.43)	-0.050* (-2.41)	-0.081* (-1.88)	-0.069* (-2.18)	0.030 (1.15)
Cayambe region	- 0.003 (0.14)	-0.023 (-1.33)	-0.116* (-2.29)	-0.026 (-1.03)	0.000 (0.01)
Migrant in Cayambe region	0.022 (1.21)	0.035 (1.62)	-0.001 (-0.02)	0.049* (1.78)	-0.021 (-0.90)
Number of observations	992	934	364	724	335
Pseudo R^2	0.9525	0.3322	0.2083	0.2340	0.2520

n.a. Not applicable.
*Significant at the 95 percent level.
[a]Only one person per household is selected for this model to control for correlation within the household.
Note: *t*-statistics are shown in parentheses.
Source: Author's calculations based on surveys conducted in May and June 1999, see text

Men's Labor Allocation

The most important result from the housework models is that Cayambe men spent more time in housework even when measured against the impacts of other variables. The wage variables that could measure a direct substitution effect had no impact on men's time in housework. On the contrary, the Cayambe location dummy, which represents an indirect wage bargaining effect, always had a positive and significant effect on men's housework (recall from equation [5] that μ, the distribution factor, affects time allocation indirectly by being a function of the wage). Wages were significant in the other models of labor allocation, such as the models for men's and women's paid labor supply and for women's unpaid labor supply, which shows that the wage variables generally had the expected effects.

For married male heads of households, the only effects that were significant were migrant status, the Cayambe dummy, and the urban dummy. The lack of significance of the other variables is assumed to stem from the loss of precision from using CLAD and the one-person-per-household sample. Despite these limitations, the results point to a strong bargaining effect on men's housework. In the CLAD model for all men, the magnitude of the impacts also shows the importance of the Cayambe region in determining men's housework. The estimated coefficient for Cayambe is about twice as large as the other significant dummy variables (married and living in an urban location).

Household-level effects were also important, showing that men's contribution to housework grew with the number of children and shrank with the total number of household members (because more members can share the work). The importance of the gender composition of the household was clearly evident: Men's housework shares declined as the ratio of women to men in the household increased. Assets were also included to test whether men in wealthier families would feel less pressure to take on housework, and this was shown to be the case.[11]

The determinants of men's paid labor supply are characteristic of classic labor supply. There is no evidence of men's paid labor being influenced by the flower sector or the gender composition of the household. Age had a positive and diminishing effect on the share of hours worked, but education had an unexpected negative effect. Own wages had a positive effect in the larger model, but wives' wages did not in the smaller one. Men who were married or had been married worked more than single men did, and men worked more if there were more children in the home. Interestingly, the Cayambe location did not have an impact on men's share of hours worked, suggesting that men had similar opportunities to work in both regions.

11. Individual assets would have been preferable, but the data proved to be too poor to be useful at the individual level. Most assets were reported at the household level.

Women's Labor Allocation

The results from the estimation of the housework model for women reveal the gender-specific nature of this work as well as the growing influence of the labor market. In the models for all women, age had a positive and diminishing effect on housework. Having higher education had no impact, but being married or previously married had a positive impact on women's housework, all as expected. The effect of wages was significant and negative, indicating that price substitution was relevant to women's unpaid labor allocation. The Cayambe dummy was significant only in the married women's housework model, but it was negative, indicating the other side of the bargaining effect: Women bargained in Cayambe to do less housework. As it did for married men, the CLAD model for women had fewer significant impacts than the other models, but household characteristics and Cayambe were the main determinants.

Household characteristics were important determinants of women's time in housework, and their effects were similar to those for men. A higher ratio of women to men in the home reduced the woman's share of household work.

The estimates of women's paid labor supply reflect gender constraints more than those of men's paid labor supply. Age had the typical human capital effect (positive and diminishing), but education did not, perhaps because so much of the available work for women was unskilled. Marital status had a negative effect on women's labor supply (the opposite of its effect on men), and being previously married had a positive impact on women's paid labor shares. A woman worked less outside the home if her husband was older than she was, and she worked less if her husband had more education than she did. Husbands' wages had no impact on women's labor supply, refuting the notion that men's labor allocation in the home reflects price substitution.

Household characteristics had the expected impacts on women's paid labor supply (the same as the effects on men), except that the ratio of females to males in the home was a positive factor (it was not for men). This effect can be understood in two ways. First, the more women there were to share housework, the freer they were to pursue outside work. Second, once one woman had broken the cultural barrier against women working outside the home, it was easier for other women in the household to work as well.

V. Conclusions

The presence of the flower industry and other household and individual factors affected time use patterns in Ecuador in many surprising ways. The most compelling evidence of the industry's impact is on men's increased participation in housework. Married men in Cayambe spent twice as much time on housework as did men in Cotocachi, and this was clearly related to women's increased participation in the labor force. Employment in the flower industry was linked to even higher levels of men's time in housework. Even households in which the

wife did not work may have been affected. In those households married men did more housework than their counterparts in Cotocachi, but the difference was not statistically significant. Women in Cayambe did less housework than did women in Cotocachi.

The econometric results provide even stronger support for the argument that women's employment in the flower industry led to increases in men's housework through a bargaining effect. The findings that men's time in housework was not directly affected by wages, or by their wives' wages if they were married, and that wages were significant in the other labor allocation models together provide strong evidence against a wage substitution effect determining men's housework. The consistently positive effect of the Cayambe location dummy on men's housework makes the case instead that higher women's wages enabled women to bargain for a redistribution of housework in which men did more than they otherwise would have.

Large gender differences in time use were found independent of the effect of the flower sector. Men worked three-fourths of the time that women did when including housework, leaving women with much less leisure time than men. This is a fundamental inequity. Women in Cayambe did not work more total time in paid and unpaid, dispelling a frequent criticism of agricultural export development that maintains that women are unduly burdened by work in the industry. Women did spend more hours working than men, but this is apparently a result of their culturally assigned housework responsibilities and not a result of the availability of employment. There may be other reasons to criticize the flower industry, but the gender impacts are arguably positive on balance given that employment for women appears to lead to cultural change. By extension the trade liberalization policies that led to the growth in this employment should be recognized as an important component in the expansion of opportunities for women. As one woman in a focus group in Cayambe said, "Once a woman starts earning an income, she can set down her own conditions because she becomes more independent."

References

Alderman, H., and S. Chishti. 1991. "Simultaneous Determination of Household and Market-Oriented Activities of Women in Rural Pakistan." *Research in Population Economics* 7:245–65.

Apps, P., and R. Rees. 1997. "Collective Labor Supply and Household Production." *Journal of Political Economy* 105(1):178–90.

Barham, B., M. Clark, E. Katz, and R. Shurman. 1992. "Nontraditional Agricultural Exports in Latin America." *Latin American Research Review* 27(2):43–82.

Bittman, M. 1999. "Parenthood without Penalty: Time Use and Public Policy in Australia and Finland." *Feminist Economics* 5(3):27–42.

Browning M., and P. A. Chiappori. 1998. "Efficient Intra-Household Allocations: A General Characterization and Empirical Tests." *Econometrica* 66(6):1241–78.

Buchinsky, Moshe. 1994. "Changes in the US Wage Structure 1963–1987: Application of Quantile Regression." *Econometrica* 62(2):405–58.

Chiappori, P. A. 1997. "Introducing Household Production in Collective Models of Labor Supply." *Journal of Political Economy* 105(1):191–209.

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore, Md.: Johns Hopkins University Press.

Fafchamps, M., and A. Quisumbing. 1998. "Social Roles, Human Capital, and the Intrahousehold Division of Labor: Evidence from Pakistan." International Food Policy Research Institute, Washington, D.C.

Ilahi, N. 1999. "Gender and the Allocation of Time and Tasks: What Have We Learned from the Empirical Literature?" World Bank, Latin America and the Caribbean Department, Washington, D.C.

Jenkins, S. P., and N. C. O'Leary. 1997. "Gender Differentials in Domestic Work, Market Work, and Total Work Time: United Kingdom Time Budget Survey Evidence for 1974/75 and 1987." *Scottish Journal of Political Economy* 2(44):153–64.

Juster, T. F., and F. P. Stafford. 1991. "The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement." *Journal of Economic Literature* 29(2):471–522.

Khandker, S. R. 1988. "Determinants of Women's Time Allocation in Rural Bangladesh." *Economic Development and Cultural Change* 37(1):111–26.

Manchester, J., and D. Stapleton. 1991. "On Measuring the Progress of Women's Quest for Economic Equality." *Journal of Human Resources* 26(3):562–79.

Newman, C., P. Larreamendy, and A. Maldonado. Forthcoming. *Mujeres y Floricultura: Cambios y Consecuencias en el Hogar*. Quito: Abya Ayala.

Powell, James L. 1984. "Least Absolute Deviations Estimation for the Censored Regression Model." *Journal of Econometrics* 25(3):303–25.

Robinson, J., and J. Gershuny. 1994. "Measuring Hours of Paid Work: Time-Diary vs. Estimate Questions." ILO *Bulletin of Labor Statistics* 1:11–17.

Quiroz, J. A., and R. A. Chumacero. 1996. "Reformas Macroeconomicas y el Sector Agricola, Ecuador: 1990–1995." In M. Whitaker, ed. *Evaluación de las Reformas a las Políticas Agricolas en el Ecuador*. Vol. II. *Estudios Detallados*.

Skoufias, E. 1993. "Labor Market Opportunities and Intrafamily Time Allocation in Rural Households in South Asia." *Journal of Development Economics* 40(2):277–310.

Thrupp, L. A. 1995. *Bittersweet Harvests for Global Supermarkets: Challenges in Latin America's Export Boom*. Washington, D.C.: World Resources Institute.

van der Lippe, T., and J. Siegers. 1994. "Division of Household and Paid Labor between Partners: Effects of Relative Wage Rates and Social Norms." *Kyklos* 47(1):109–36.

World Bank. 1996. Uganda: *The Challenge of Growth and Poverty Reduction*. Washington, D.C.

———. 2001. *Engendering Development through Gender Equality in Rights, Resources, and Voice*. New York: Oxford University Press.

THE WORLD BANK ECONOMIC REVIEW, VOL. 16, NO. 3 397–423

The Distributional Impacts of Indonesia's Financial Crisis on Household Welfare: A "Rapid Response" Methodology

Jed Friedman and James Levinsohn

Analyzing the distributional impacts of economic crises is an ever more pressing need. If policymakers are to intervene to help those most adversely affected, they need to identify those who have been hurt most and estimate the magnitude of the harm they have suffered. They must also respond in a timely manner. This article develops a simple methodology for measuring these effects and applies it to analyze the impact of the Indonesian economic crisis on household welfare. Using only pre-crisis household information, it estimates the compensating variation for Indonesian households following the 1997 Asian currency crisis and then explores the results with flexible nonparametric methods. It finds that virtually every household was severely affected, although the urban poor fared the worst. The ability of poor rural households to produce food mitigated the worst consequences of the high inflation. The distributional consequences are the same whether or not households are permitted to substitute toward relatively cheaper goods. The geographic location of the household matters even within urban or rural areas and household income categories. Households with young children may have suffered disproportionately large adverse effects.

The collapse of the Indonesian rupiah during the 1997 Asian currency crisis precipitated a 12 percent decline in Indonesia's gross domestic product the following year as well as rampant inflation. In an 18-month span, food prices nearly tripled, and prices for other goods also rose substantially. The degree to which Indonesian households were vulnerable to these changes depended on a mix of factors, including the types of goods the household consumed, which goods' prices rose fastest, and the degree to which changes in income were able to buffer households from the brunt of the price shocks.

In this article we focus on the first two factors—household consumption choices and goods price changes—to explore how the price changes affected households

Jed Friedman is Associate Economist with the RAND Corporation, Santa Monica, California. His e-mail address is jedf@rand.org. James Levinsohn is a Professor at the Ford School of Public Policy and the Department of Economics at the University of Michigan, Ann Arbor, and a Research Associate of the National Bureau of Economic Research. His e-mail address is jamesl@umich edu The authors are grateful to the World Bank for partially funding this research. Alan Winters, Will Martin, and three anonymous referees provided helpful comments. Special thanks are due to Angus Deaton, who essentially suggested the methods used in this article.

DOI 10.1093/wber/lhf001

across income levels and regions of Indonesia. We develop and apply a simple methodology that requires only the sort of data that are often readily available immediately after an economic crisis. Although we focus on the Indonesian experience, our methodology is intended to be applicable to a wide range of situations and countries. We hope to present a meaningful and straightforward methodology that can be adopted to analyze the distributional consequences of financial crises and inflation anywhere in the world.

A careful and definitive investigation of the impacts of the Indonesian currency crisis and potential differential impacts across levels of living ideally requires detailed income and expenditure information for a large nationally representative sample of households, both before and after the crisis. These data do not exist. Some sources of data, in particular the Indonesian Family Life Survey (IFLS), approach this ideal. Waves 2 and 2+ of the IFLS gathered pre- and postcrisis information on a panel of 2,500 households. Studies using these data have investigated the impacts of the crisis on consumption, employment, and education (see Frankenberg, Thomas, and Beegle 1999, Smith and others 2002, Thomas and others 2001). Although these data present an impressive depth of information for households and communities, they are limited to a relatively small sample of households in a minority of provinces and thus cannot speak to the breadth of the crisis across the sweep of Indonesian geography or income distribution. About a year and a half after the release of the initial IFLS-based reports (and three years after the onset of the crisis), studies employing nationally representative postcrisis (nonpanel) household information have begun to appear, an example of which is Suryahadi and others (2000).

Our approach is distinct from these studies in that we use only household data collected before the onset of the crisis. We then match these consumption data with information on commodity price changes brought on by the crisis to calculate simple measures of compensating variation—the amount of money sufficient to compensate households following price changes and enable a return to precrisis levels of utility. We calculate this compensating variation with a variety of methods and compare and contrast the strengths and weaknesses of each approach. Our analysis employs data with sufficient degrees of freedom to allow an exploration of differences in compensating variation across the spectrum of household income and location.

Because the analysis presented here requires only pre-crisis household consumption information and price change data, our approach is applicable to other national and subnational settings. Many countries now conduct periodic household consumption surveys, and even more collect more or less current price data used for computing price indices. An important benefit of these methods is the immediacy of the findings. Postcrisis household surveys can yield valuable and even definitive information, but these data are available only after the substantial lag needed for data collection and processing. In the face of rapid economic change and social disruption, the information needs of policymakers are immediate. We propose a rapid response method that can be implemented at the onset

of a financial crisis, well before postcrisis household data can be collected and disseminated.

I. The Data

We match household-level data on consumption with province-level information on commodity price changes. The consumption data come from the consumption module of the 1996 National Socio-Economic Survey, known by the Indonesian acronym SUSENAS. Indonesia conducts this extensive household consumption survey every three years; the 1996 wave, which surveyed 61,965 households, was the most recent survey before the onset of the crisis. These surveys tend to be large, but they are not panels, that is, there is no systematic effort to track the same households over time. They do cover the entire geographic range of the country and contain very detailed consumption data on 306 food and nonfood goods. SUSENAS also records whether food goods were purchased in the market or produced by the household. If food is self-produced, SUSENAS imputes a value at prevailing local prices. SUSENAS also imputes a rental value for owned housing.

SUSENAS does not contain information on prices. Rather, the data enable the computation of unit values, defined as the expenditure for a particular good divided by the quantity consumed. These unit values may differ across households that face identical prices due to differences in the choice of consumption quality. For example, though all households in a village may face the same prices for high-quality and low-quality rice, the unit values recorded for a household that bought mostly high-quality rice will be higher than the unit values recorded for the household that bought mostly low-quality rice. These higher unit values reflect the higher mean quality of total rice purchases. Demand systems can be estimated with unit values in lieu of actual prices by exploiting the spatial variation in the data using methods developed by Deaton (1988, 1990, 1997). The unit value data and demand system estimation technique are used in subsequent sections.

We also have recent price data supplied by the Indonesian Central Statistical Office (the Biro Pusat Statistik, or BPS). These data contain monthly price observations for 44 cities throughout the country from January 1997 to October 1998. This period, which begins before the advent of the crisis, spans the steep devaluation of the rupiah and subsequent (and temporary) stabilization at the new higher rate. We employ a single price change measure, the percentage change in price from January 1997 to October 1998. By adopting such a long time period from before the onset of rapid inflation until after the inflation had largely abated, we hope to capture a robust measure of the price changes associated with the crisis.

The price data provide information on both aggregate goods, such as food and housing, and individual goods, such as cassava and petrol. There are about 700 goods with observed prices in the data. However, the types of goods ob-

served vary by city, perhaps reflecting taste and consumption heterogeneity throughout the country. On average, a city has price information on about 350 goods. Jakarta has as many as 440 goods listed, whereas some small cities have price information for only 300 goods.

Each of the 27 Indonesian provinces is represented by at least one city in the price data. To match households from the SUSENAS data to as local a price change as possible, we calculate province-specific price changes from the city-level data. For provinces with only one provincial city in the price data, we take those price changes as representative of the entire province. For provinces with more than one city in the price data, we calculate an average provincial price change using city-specific 1996 population weights.

The accuracy of this extrapolation of city price data to an entire province will surely vary with the size and characteristics of the province considered. For example, Jakarta, the national capital, is its own province, and the observed price changes will fairly accurately represent the price changes faced by residents throughout the province. In contrast, the price changes for Irian Jaya, a vast mountainous province, are based on price changes observed in the provincial capital, Jayapura. Those price changes may not be a completely accurate proxy for price changes in remote rural areas. Indeed, Frankenberg, Thomas, and Beegle (1999) suggest that postcrisis inflation in rural areas may have been 5 percent higher than in urban areas. We frequently report separate results for rural and urban households; the fact that price data were collected in cities should be kept in mind as those results are reviewed.[1]

We match the price change data with the consumption data to calculate the measures of compensating variation, detailed in the next section. There are 219 products and product aggregates that appear in both the SUSENAS and our price data. We attempt to match goods across the two data sets at the lowest level of aggregation possible. For the case of food (both raw and prepared) we were able to match 155 individual goods between the two data sets. In the case of nonfood items we matched 64 different goods, both individual goods, such as firewood and kerosene, and aggregate goods, such as toiletries and men's clothing.

1. We explored how unit values in urban and rural areas have changed across different SUSENAS survey years (1984, 1987, 1990, 1993, and 1996) to investigate potential differences in the time trend of urban and rural prices. We accomplished this by fixing a specific food bundle and then pricing this bundle separately for both urban and rural areas in each survey year. Instead of actual prices, however, we used the mean national urban or rural unit values as our price measure. In essence, we generated separate series of price indices for urban and rural areas The time trends of these indices are virtually identical. For example, the urban index increased 187 percent over the 1984–96 period, and the rural index increased 182 percent. Each three-year change in urban unit values is even closer in magnitude to its rural counterpart. This pre-crisis comovement of urban and rural unit values suggests that urban and rural prices may behave in a similar manner following the crisis and thus our extension of urban price changes to rural areas may not introduce significant bias.

For certain groups of goods, the price data are more disaggregated than the consumption data reported in SUSENAS. In these cases, we take the simple average of the underlying price changes and apply it to the larger consumption aggregate. In other cases we also aggregate commodity expenditure categories in SUSENAS to match a larger product category in the price data. The match between the price data and the consumption data is good but not perfect; we have detailed price data for most (but not all) of the goods that make up a household's total expenditure. On average, expenditures on the matched goods account for 79 percent of a household's total expenditure; this proportion is slightly greater for poor households and slightly less for wealthy ones.

For use in subsequent analysis, we calculate the budget shares of each of the 219 products and product aggregates based on the reported monthly expenditures for each item. For durable goods and other nonfood items, we use the monthly average of annual expenditure, and not the expenditures in the month preceding the survey, to more accurately measure monthly expenditures for durables that are infrequently purchased. Table 1 gives an overview of the consumption data by reporting budget shares for selected composite goods. These goods are not chosen from among the 219 items but rather are composite aggregates constructed only for the expositional purposes of table 1. Even the rice good in the first row of table 1 is an aggregate of three different varieties. To highlight the heterogeneity in consumption patterns, we report mean budget shares for the entire sample as well as for the top and bottom decile of household expenditures. Clearly, rice is the single most important commodity, as measured by the budget share, for the majority of Indonesians. Households in the bottom expenditure decile devote more than a quarter of all outlays to rice, whereas for the mean household a still substantial 16 percent of total expenditures goes toward rice. The next most important aggregate consumption category encompasses housing and utilities, especially so for the top expenditure decile, where 22 percent of spending goes toward those ends.

Alongside the budget shares, table 1 also reports the average price increase for each product aggregate. This is accomplished by calculating the household-specific price increase of the composite goods using household expenditure shares to weight the price increases of each constituent individual good. We then average these household-specific price increases over all households.

By any measure, the inflationary impacts of the crisis were large. The all-important rice price increased an average of almost 200 percent, and the prices for many foodstuffs increased more than 100 percent. Nonfood prices did not rise nearly as rapidly, with the housing and utilities price increasing least (only 24 percent on average). Listed next to the mean price increases are the standard deviations of the price increases for the aggregate goods. Due to the constructed nature of the reported price changes, variations in price change will arise due to both geographic variation in price changes and household variation in consumption of individual goods. For rice, a relatively homogenous good, all of the varia-

TABLE 1. Budget Shares and Price Changes for Selected Aggregate Goods

Product aggregate	Mean budget shares			Price changes	
	Bottom decile	All households	Top decile	Mean price increase (percent)	Standard deviation
Rice	0.269	0.164	0.048	195.2	29.2
Other cereals and tubers	0.030	0.010	0.003	137.5	101.8
Fish	0.033	0.040	0.032	89.1	67.4
Meat	0.008	0.025	0.040	97.0	49.3
Dairy and eggs	0.015	0.027	0.031	117.1	31.9
Vegetables	0.034	0.032	0 020	200.3	129.5
Pulses, tofu, & tempeh	0.025	0.023	0.012	95.2	76.0
Fruit	0.016	0.021	0.027	103.7	61.3
Oils	0.040	0.030	0.015	122.0	74.8
Sugar, coffee, & tea	0.041	0.034	0.019	142.9	28.3
Prepared food & beverages	0.025	0 047	0.058	81.4	51.7
Alcohol, tobacco, & betel	0.039	0.049	0.031	93.9	43.8
Housing, fuel, lighting, & water	0.146	0.162	0.223	23.8	10.9
Health	0.010	0.014	0.021	50.7	32.9
Education	0.013	0.021	0.037	55.3	31.9
Clothing	0.044	0.045	0.041	84.4	25.2
Durable goods	0.013	0.034	0.075	114.3	34.3

Note: Price increases are from January 1997 through October 1998. Mean price increases are computed as the average across all households reporting positive consumption for a given good. Mean budget shares are reported for the entire sample, as well as separately for the top and bottom expenditure decile.
Source: Authors' calculations from 1996 SUSENAS and BPS Price Data.

tion in the price increase is geographical, and a standard deviation of 30 percent shows how varied the price increases actually were.[2] If the price changes for rice were distributed in a roughly normal fashion, fully one-third of households experienced an increase in the rice price outside the interval (165 percent, 225 percent). Other reported price changes combine variation in household consumption choice with regional variation in price changes; the standard deviations of these price changes tend to be larger.

Given the wide dispersion of price changes both within and across product aggregates, what a household consumes and where a household lives will go a long way toward determining the particular impacts of the crisis. The next section discusses how we measure these household-specific consequences.

II. METHODOLOGY

To consider the impacts of the price increases on household welfare, we look at changes in consumer surplus brought about by the change in prices.[3] We start

2. Although there are three rice varieties from the SUSENAS consumption data, the BPS price data provide only one price change for all rice varieties.

3. Ravallion and van de Walle (1991) adopt a conceptually similar approach in their investigation of the impact of food price policy reforms on poor Indonesian households

with a minimum expenditure function, $C(u,p)$, which, given existing prices p, relates the minimum cost needed to attain utility level u. (See Deaton and Muellbauer 1980, pp. 25–59, for a discussion of the general properties of cost functions.) A first-order Taylor expansion of the minimum expenditure function with respect to price will yield an approximation of the income required to compensate the household after a price change and to restore that household to the prechange utility level. Thus this expression will approximate the compensating variation. Noting that the partial derivative of the minimum expenditure function with respect to price yields quantities consumed, we derive this simple expression:

$$\Delta C \approx q\Delta p \tag{1}$$

where q is a $1 \times n$ vector of consumption goods quantities, Δp a $1 \times n$ vector of price changes, and n the number of consumption goods in the total demand system. We note that this first approximation of compensating variation requires information only on pre-crisis consumption quantities and on price changes; neither price levels nor, more important, postcrisis consumption choices are needed.

It is straightforward to reformulate expression (1) in terms of budget shares, w, and proportionate price changes with the following expression:

$$\Delta \ln C^h \approx \sum_{i=1}^{n} w_i^h \Delta \ln p_i^h \tag{2}$$

where i refers to individual goods in the commodity system and h refers to the household. The budget share w is simply the household cost of good i divided by pre-crisis total household expenditures. Made clear by expression (2) is the fact that any differential distributional impact of the price changes must derive both from the presence of large relative price changes and large differences in the budget shares across households. Table 1 shows that this combination of factors existed in Indonesia following the crisis.

In general, the costs of attaining pre-crisis utility levels will increase less rapidly than expression (2) may suggest, because households can substitute away from goods whose prices have risen disproportionately. Thus expression (2) provides a maximum bound on the impact of the crisis because it does not take into account the substitution toward relatively less costly products that will take place. Given the large relative price changes following the crisis, this substitution surely occurred to some extent. Expression (2) may therefore not be an entirely accurate approximation. Returning to the minimum expenditure function, a second-order Taylor expansion of the minimum expenditure function does allow for substitution behavior:

$$\Delta C \approx q\Delta p + \frac{1}{2}\Delta p^T s \Delta p. \tag{3}$$

In expression (3) q and Δp are quantity and price change vectors as before and s is the $n \times n$ matrix of compensated derivatives of demand. As we did in expres-

sion (2), we can reformulate expression (3) in terms of budget shares and proportional price changes as:

$$\Delta \ln C^h \approx \sum_{i=1}^{n} w_i^h \Delta \ln p_i^h + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \Delta \ln p_i^h \Delta \ln p_j^h \tag{4}$$

where the expression c_{ij} contains the Slutsky derivatives s_{ij} and is defined by the expression

$$c_{ij} = p_i s_{ij} p_j \,/\, C^h.$$

With some simple algebraic manipulation we can show the c_{ij} term to be equivalent to $w_i \varepsilon_{ij}$,

$$c_{ij} = \frac{p_i s_{ij} p_j}{C^h}$$

where ε_{ij} is defined as the compensated price elasticity of good i with respect to price change j. Thus we can restate expression (4) as:

$$\Delta \ln C^h \approx \sum_{i=1}^{n} w_i^h \Delta \ln p_i^h + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_i^h \varepsilon_{ij} \Delta \ln p_i^h \Delta \ln p_j^h. \tag{5}$$

We will use the two formulations of compensating variation given in expressions (2) and (5) to explore the possible differential impacts of the Indonesian currency crisis. The only additional pieces of information required in expression (5) and not found in expression (2) are the ε_{ij} terms. Thus an approximation to the compensating variation that also wishes to account for potential household substitution behavior requires estimates of a complete set of price elasticities in addition to the pre-crisis consumption quantities and postcrisis price changes.

Exactly how these elasticities are estimated depends on the types of data used in the analysis. Our task appears difficult because we have neither information on household consumption changes over time nor information on price levels at the time of the household consumption survey. Instead of prices we have unit value data. In a series of articles, Deaton (1988, 1990, 1997) presents an approach to elasticity estimation using unit values as price proxies and only a single cross-section of household information. Crucial to this approach is the recognition that prices for equivalent goods can vary greatly across space in a developing economy setting and that household survey information is often gathered in clusters to reduce survey costs. Given these insights as well as certain assumptions on how households choose the quality of goods they purchase, the clustered nature of these data can be exploited to purge the unit value data of quality components. The cross-spatial variation in these purged unit values can then be used to identify own-price or cross-price elasticities. This is the method adapted here to estimate the ε_{ij} terms.

We now summarize this method in a bit more detail before moving on. Deaton suggests adopting the following econometric specifications for the log quantity ($\ln q$) and log unit value ($\ln v$) of a particular good:

$$\ln q_{hc} = \alpha^0 + \beta^0 \ln x_{hc} + \gamma^0 z_{hc} + \varepsilon_p \ln \pi_c + f_c + u^0_{hc}$$
$$\ln v_{hc} = \alpha^1 + \beta^1 \ln x_{hc} + \gamma^1 z_{hc} + \psi \ln \pi_c + u^1_{hc}$$

where h and c index household and cluster, x represents total household expenditures, z household demographic characteristics, and π the (unobserved) price of the good. The quantity equation also contains a cluster fixed effect, f_c, and the coefficient of interest is ε_p, the price elasticity. The simplified process to be described here concerns only the estimation of own-price elasticities; cross-price terms can be added through a relatively straightforward extension. The final estimate of ε_p derives from two main steps. In the first step, the within-cluster variation of household income and other characteristics is used to estimate β and γ (because prices are constant within clusters, these parameters can be consistently estimated). The estimated coefficients are then employed to generate two variables:

$$\hat{y}^0_{hc} = \ln q_{hc} - \hat{\beta}^0 \ln x_{hc} - \hat{y}^0 z_{hc}$$
$$\hat{y}^1_{hc} = \ln v_{hc} - \hat{\beta}^1 \ln x_{hc} - \hat{y}^1 z_{hc}.$$

The next step is to calculate the cluster-level averages of y^0 and y^1. Then a "regression" of the cluster averaged y^0 on the cluster averaged y^1 will yield an estimate of the ratio of ε_p to ψ:

$$\frac{\varepsilon_p}{\Psi} = \frac{\operatorname{cov}\left(\hat{y}^0_c \hat{y}^1_c\right)}{\operatorname{var}\left(\hat{y}^1_c\right)}.$$

Combining this expression with an estimate of ψ—identified from previously estimated coefficients in an expression (not shown) determined by the model of household quality choice—enables the researcher to calculate the price elasticity estimate.[4]

If one wishes to estimate a demand system of our dimensions, some product aggregation is necessary. There are simply not enough degrees of freedom in the SUSENAS data to estimate a demand system for 219 products complete with the all-important cross-price elasticities. The types of goods for which we can estimate price elasticities are also limited by the fact that SUSENAS reports unit values solely for food goods. Hence we reduce the dimensions of the problem through aggregation and estimate elasticities for 22 composite goods—21 aggregate food goods and a residual nonfood consumption category. A subsequent table in the next section lists each of these aggregate goods.

Another issue concerns the services provided by owner-occupied housing and self-produced agriculture. Many households, especially in rural areas, own their own home. Although the price of housing has increased, these households are not any better or worse off in an absolute sense (they are still living in the same house). However, these households are better off relative to those who do not

4. This brief discussion ignores the identification of ψ as well as the important role of corrections for measurement error found in the original series of articles. We refer readers to those works for more extensive presentation and discussion.

own their own homes. We choose to account for these services provided by owner-occupied housing by treating the imputed rental value for these homes as a negative expenditure.

Many households, mostly rural, also produce some of their own food. Households that consume self-produced foodstuffs are also potential net exporters of agricultural products. As the price of food rose, the value of their production also increased. Clearly, if the household were a net exporter of food, the household would benefit from the price increase. To the extent that a household produced some of its own food, such production would mute the impact of price increases relative to a household that purchased food in the market. We account for self-produced agricultural products by treating the imputed value of self-produced food as a negative expenditure.[5]

Once the budget share and price change data have been matched and the price elasticities estimated, we calculate our two measures of compensating variation for each household. So that we can explore in a flexible manner how expressions (2) and (5) vary across levels of living, our principal approach is nonparametric. Specifically, we use locally weighted least squares to estimate the compensating variation at each point in the income distribution (see Fan 1992 for an introduction of this method). Local observations were weighted with a biweight kernel. After experimentation, we choose to adopt a bandwidth of 0.4 units of the independent variable (log per capita monthly household expenditures).

We use two measures to assess a household's level of living. The first is per capita household expenditure; the second is a binary poor/nonpoor measure dependent on whether the household's per capita expenditure exceeds or falls below a predetermined poverty line.

Perhaps the most standard approach to measuring the level-of-living in a developing economy setting is to use some estimate of household expenditures. In this view, the level of household consumption constitutes the lion's share of total household utility, and total consumption is most easily proxied by the household's actual expenditures. Expenditure levels are generally viewed as a better measure of welfare than income because the ability to smooth consumption in the presence of income shocks suggests that expenditures (rather than income) more closely track actual welfare.[6]

5. This approach will understate the effects of food price increases to the extent that we do not observe or adjust for price increases of intermediate inputs used in agricultural production. There is a long-standing debate over whether shadow prices in rural households engaged in agricultural production equate market prices for agricultural inputs such as labor or land. To the extent that these shadow prices diverge from market prices, the "valuation" for self-produced food, based on market prices, will not be entirely accurate. Benjamin (1992) presents evidence from rural Java that household shadow prices for agricultural inputs such as labor are not significantly different from market prices.

6. Chaudhuri and Ravallion (1994) investigate the competing merits of using these two welfare indicators and find little difference when the goal is to distinguish poor from nonpoor households. This article remains within the standard literature and uses household expenditures as a main measure of household welfare.

In addition to this continuous measure of level of living, an alternative binary poverty measure is adopted. A household is deemed poor if its per capita expenditure lies below a fixed poverty line. The poverty lines used here are calculated from the 1996 SUSENAS using a cost of basic needs approach to poverty determination, as set forth in Bidani and Ravallion (1993), Ravallion (1994), and Ravallion and Bidani (1994); the details of the particular method used here are presented in Friedman (forthcoming). The general approach is summarized as follows. A nutritionally adequate food bundle (with nutritional guidelines stipulated by WHO/FAO/UNU 1985) that reflects the actual consumption choices of Indonesian households is determined and then priced. The total cost of this bundle is scaled upward by an econometrically estimated factor that represents the cost of essential nonfood goods. This final value, which we take as the poverty line, proxies the total cost of essential food and nonfood consumption needs. Due to important differences in relative prices between urban and rural areas, poverty lines are computed separately for each area. For the 1996 SUSENAS, this method translates into a poverty line of 36,956 rupiahs per person per month in urban areas and 32,521 rupiahs in rural areas. These values yield poverty headcounts of 9.3 percent in urban areas and 24.9 percent in rural areas.

III. Results

The impacts of the crisis were probably not uniform. Instead, household consumption choices, sources of income, and location mattered greatly in determining the specific impact. The diversity of impacts was due both to wide geographical variation in price changes and wide variation in household structure and consumption. An earlier article (Levinsohn and others forthcoming) explores this heterogeneity in detail. Our focus here is solely on the relative differences in the compensating variation measures across the income distribution. This relative difference is exhibited clearly in table 2, which reports summary mean values of expression (2) by decile of household expenditure as well as poor/nonpoor status. When looking at all households, we see that the compensating variation has an inverted U-shape, with the lowest decile having an average compensating variation of 73 percent of initial household expenditures, rising to a 85 percent of household expenditures for those in the sixth, seventh, and eighth deciles and falling back to 77 percent for households in the top decile. From this perspective, it was the Indonesian households in the middle of the distribution that were most adversely affected by the price changes. Indeed, poor households needed to earn less income (as a proportion of initial expenditures) than nonpoor households—77 percent versus 82 percent—to return to original consumption levels.

However, we see in the next two columns that this story obscures important differences between households in rural and urban areas. For urban areas, households in the lower deciles needed the greatest relative amount of new income to return to pre-crisis consumption levels, and indeed this amount declined monotonically as household expenditures increase. For rural areas, lower-income

TABLE 2. Compensating Variation by Expenditure Decile and Poor/Nonpoor Status

Expenditure decile	All households	Urban	Rural
1	0.73	1.08	0.67
2	0.79	1.03	0.73
3	0.82	1.00	0.74
4	0.83	0.96	0.77
5	0.84	0.93	0.77
6	0.85	0.92	0.78
7	0.85	0.89	0.78
8	0.85	0.84	0.79
9	0.84	0.81	0.79
1	0.77	0.70	0.81
Poor	0.77	1.09	0.70
Nonpoor	0.82	0.90	0.78
All households	0.82	0.91	0.76

Note: Compensating variation measured as a proportion of 1996 household expenditures.

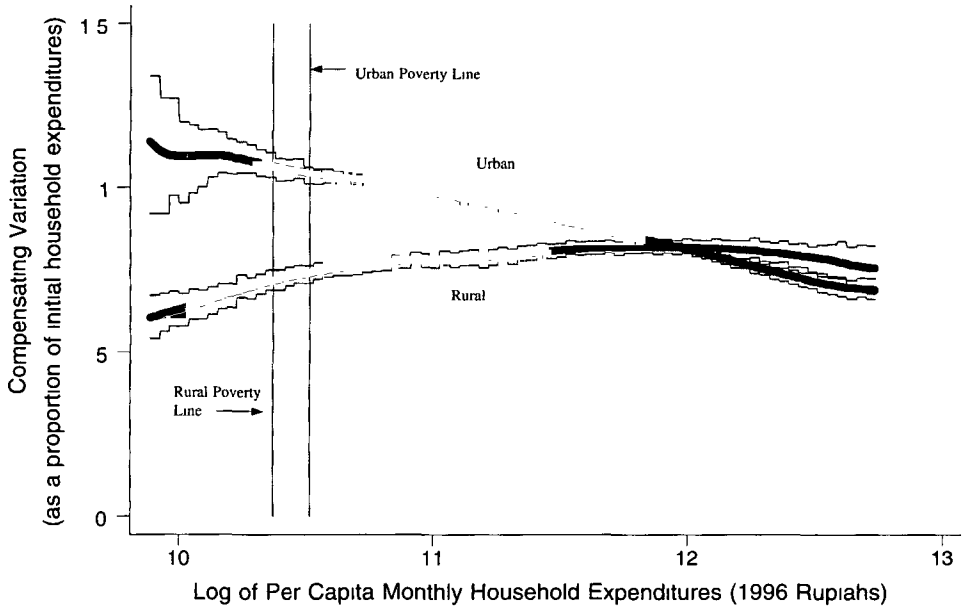
Source: Authors' calculations from 1996 SUSENAS and BPS Price Data.

households needed the least relative compensation, and this amount increases monotonically with expenditures. Table 2 suggests that it is the urban poor who were the most adversely affected by the crisis needing, on average, 109 percent of their pre-crisis income to reach pre-crisis utility levels. In contrast, the rural poor required only 70 percent of their pre-crisis income. In general, urban households, composed mainly of households that do not grow their own food, fared the worst under the price changes.

The same story is captured in figure 1, which depicts the entire distribution of the compensating variation measure as estimated by locally weighted least squares. The figure also includes the urban and rural poverty lines for reference, as well as bootstrapped 95 percent confidence intervals for each regression line.[7] The urban regression line declines almost entirely monotonically from its peak at the bottom of the income distribution to its trough at the high end of the distribution. In contrast, the rural regression line rises from its low at the bottom of the distribution and then flattens out for households beyond the top third of the distribution. After this point in the expenditure distribution there are virtually no differences in the compensating variation measure and no statistically significant difference between urban and rural households. However, the large differences between poor urban and poor rural households are indeed significant at conventional levels. As figure 1 shows clearly, the urban poor were the most adversely affected, and the rural poor were perhaps the least affected.

7. The bootstrapped standard errors were estimated with 50 draws (with replacement) from the total sample and took into account the clustered nature of the underlying survey data.

FIGURE 1. Compensating Variation, with 95 Percent Confidence Intervals



The results in table 2 and figure 1 derive from expression (2) and likely overstate the true compensating variation because expression (2) does not allow for the substitution behavior that surely occurred to some degree. As already discussed, the addition of the second-order terms in expression (5) may give a better approximation to the true compensating variation because it includes substitution terms. These elasticities were identified by the spatial variation of consumption choices and unit values in the 1996 SUSENAS using the methods discussed earlier. Before moving on to estimates of expression (5), table 3 presents these estimated price elasticities for the 22 composite good demand system (21 food goods and the residual nonfood category). The own-price elasticities for each composite good are located on the diagonals in the price matrix, and they are negative for almost every good. The own-price elasticity of rice is estimated to be –0.48, exactly equal to that found by Case (1991) using earlier SUSENAS data and different methods of estimation. The three goods (preserved meat, prepared beverages, and alcohol) that are estimated to have positive own-price elasticities are goods that have substantially fewer positive consumption values than the other goods. In other words, these goods are not widely consumed; as such their elasticities are not likely to be precisely estimated. The cross-price elasticities are generally smaller in magnitude than the own-price elasticities and, of course, vary in sign depending on whether the data suggest a particular pair of goods to be either substitutes or complements.

With this matrix of own- and cross-price elasticities we reestimate the compensating variation using expression (5) and then contrast the results with those

TABLE 3. Estimated Price Elasticities for Aggregate Food Goods and Residual Consumption

Product	Rice	Other cereals	Tubers	Fresh Fish	Preserved Fish	Fresh meat	Preserved meat	Eggs	Dairy	Green vegetables
Rice	-0.479	0 082	-0 032	-0 029	-0 038	0 098	-0 016	-0 008	-0 009	-0 018
Other cereals	2 762	-5 046	-0 413	-0 074	0 387	-0 200	-0 134	-0 300	-0 014	0 048
Tubers	2 521	-0 127	-0 590	0 233	0 205	-0 672	0 087	-0 531	-0.167	-0 919
Fresh fish	-0 383	0.027	0 217	-0 996	0 026	0 169	0 219	-0 087	-0.012	0 026
Preserved fish	-0 533	-0 295	-0 059	0 373	-0 686	0 013	-0.015	-0 022	0 138	-0 103
Fresh meat	0 042	0 073	-0 046	0 056	0 118	-0 616	-0 004	-0 134	0 109	-0 135
Preserved meat	-0 224	0 318	0 127	0.256	0.254	-0.418	0.955	-0 281	-0.260	-0 215
Eggs	-0 458	0.128	0 013	-0 006	-0.080	0 084	-0 080	-0 985	-0 028	0 113
Dairy	-0 194	0 121	0 097	-0 072	-0.083	-0.216	0.548	0 040	-0 133	0 077
Green vegetables	-0 384	0.097	0 189	-0.202	-0.041	-0 067	0 136	0 014	-0 023	-0.789
Other vegetables	-0 465	-0 005	-0 042	0 125	0.017	-0.115	0.074	0 034	-0 004	0 002
Pulses	-0 406	0.367	-0 001	-0 153	-0 064	0 266	-0 271	-0 248	-0 474	-0 014
Tofu & tempeh	-0.104	0 077	0 010	0 102	-0.033	-0.111	-0.159	0 160	-0 025	0 017
Fruit	-0 181	-0.144	-0 141	0 098	-0 006	-0 253	0 044	-0 147	-0 110	-0 021
Oils	-0.238	-0.012	0 027	-0 143	-0 003	-0 136	-0 019	-0 004	0 007	-0 009
Beverage additives	-0 173	0 059	0 044	-0 167	0 013	0.001	-0 111	-0 047	-0 106	0 064
Spices	-0.210	-0 018	0 104	-0 072	-0.007	0.000	-0 034	-0.057	-0.107	0 032
Other food	0 140	-0 056	0.069	-0 027	0 004	-0 238	0 098	0.112	0 013	0 029
Prepared food	0 020	0.243	0 055	0.092	-0 006	-0 037	-0 037	0 060	-0 093	0 042
Prepared beverages	-0 429	0 026	-0 083	0 246	0 005	0 034	0.259	-0 191	-0.203	0 146
Alcohol	-2 806	-0 681	0 161	0 859	0 265	-2 175	2 039	0 506	-0 012	-0 770
Tobacco & betel	-0.441	0 053	0.001	-0 104	-0.037	0 151	-0 182	0 001	-0 025	0 033
Other consumption	0 010	0 017	0 008	-0.010	-0 003	0.019	-0 003	-0 013	-0.008	-0 002

Source: Authors' calculations from 1996 SUSENAS.

TABLE 3. (*continued*)

Other vegetables	Pulses	Tofu and tempah	Fruit	Oils	Beverage Additives	Spices	Other food	Prepared food	Prepared beverages	Alcohol	Tobacco and betel	Other consumption
0 003	0 001	0 023	-0 053	0 058	0 032	-0 042	-0 037	-0 026	0 032	0 036	-0 138	0 274
-0 667	0 002	0 262	-0 195	0 289	-0 239	0 086	-0 210	-0 292	0 430	0 649	-0.276	2 684
-0 147	0 010	0 467	-0 142	-0.590	-0 387	-0 196	0 621	0 225	0 660	-0 292	0 052	-0 822
0 071	0 042	-0.038	0 165	-0 012	-0 104	0 135	-0 003	-0 025	0 185	0 128	0 304	-1 065
0 091	-0 186	0 524	-0 198	0 011	-0 014	-0 093	0 058	-0 043	-0 143	0.164	-0 603	1 066
-0 173	0 092	-0 091	-0 023	-0 182	0 052	0 065	0 113	0 163	-0 075	-0 047	-0 063	-1 091
-0 106	0 259	-0 212	-0 094	-0 276	0 325	0 007	0 372	0 269	1 018	0 330	0 540	-5 594
0 031	0 071	-0 052	0 027	0.019	-0 155	0 131	0 033	0 007	-0 019	-0 135	0.015	0 390
-0 002	0 122	-0 097	0 132	-0.150	0 216	0 283	-0 145	0 124	-0 296	-0 037	0 441	-2 696
0 057	0 060	-0 099	0 033	0 110	0 106	0.138	0 004	0 064	-0 023	0.257	-0 086	-0 088
-0 840	0 041	-0 001	-0 114	-0 134	0 031	0 061	-0 095	-0 028	0 156	0 242	-0 096	0 432
0 169	-0 772	-0.136	0.120	-0 135	-0.378	0 216	0 192	-0 048	-0.149	-0 281	0 038	0 936
0 052	-0 057	-0 965	0 035	0 032	0 059	0 292	-0 194	0 022	-0 056	-0 106	-0 069	0 606
0 003	0 115	-0 006	-0 831	-0 111	-0.094	0 074	0.005	0 030	-0.060	-0 069	0 016	0 478
0 022	0 008	0 058	-0 064	-1 003	-0 039	0 038	-0 017	-0 031	0 144	0 019	0 036	0 789
-0.036	0 056	-0 029	-0 082	0.109	-0 625	0 017	0 029	-0 013	0 132	0 101	0 019	0 055
0 023	0 049	-0 073	-0 016	0 032	-0 033	-0 305	0 027	-0 062	-0 005	-0 080	-0 114	0 248
-0 010	0 091	-0 059	0 027	0 019	0 039	0 177	-1 161	0 064	-0 008	0 114	0 139	-0 763
-0 033	-0 078	-0 090	0 209	-0 044	-0.060	0 062	0 089	-0 775	-0 339	-0.135	0 124	-0 401
-0 227	-0 040	-0 206	0.025	-0 162	0 720	0 137	-0.231	0 310	1 912	-0 517	-0 296	-3 857
0 447	0 143	0.204	-0 311	-0 545	-0 602	0 668	-0 569	1 183	1 501	6 106	1 226	-9 039
0 030	-0 070	0 061	-0 017	0.070	0.067	-0 025	0 026	-0.147	-0 260	0 023	-0 876	0 664
0 002	0 007	0 005	-0 001	0 002	0 013	0.029	-0 003	-0 002	-0 001	0 020	-0.011	-0 482

obtained using expression (2). The comparisons, estimated with locally weighted least squares, are separated by urban and rural household location (figure 2). As is readily apparent, the qualitative conclusions drawn with expression (2) also hold with results that now allow for substitution behavior. Across urban areas the compensating variation declines as household expenditures increase, again suggesting that poor urban households are affected the most severely by the price changes. Similarly, poor rural households appear to fare the best, with little difference between wealthier urban and rural households.[8]

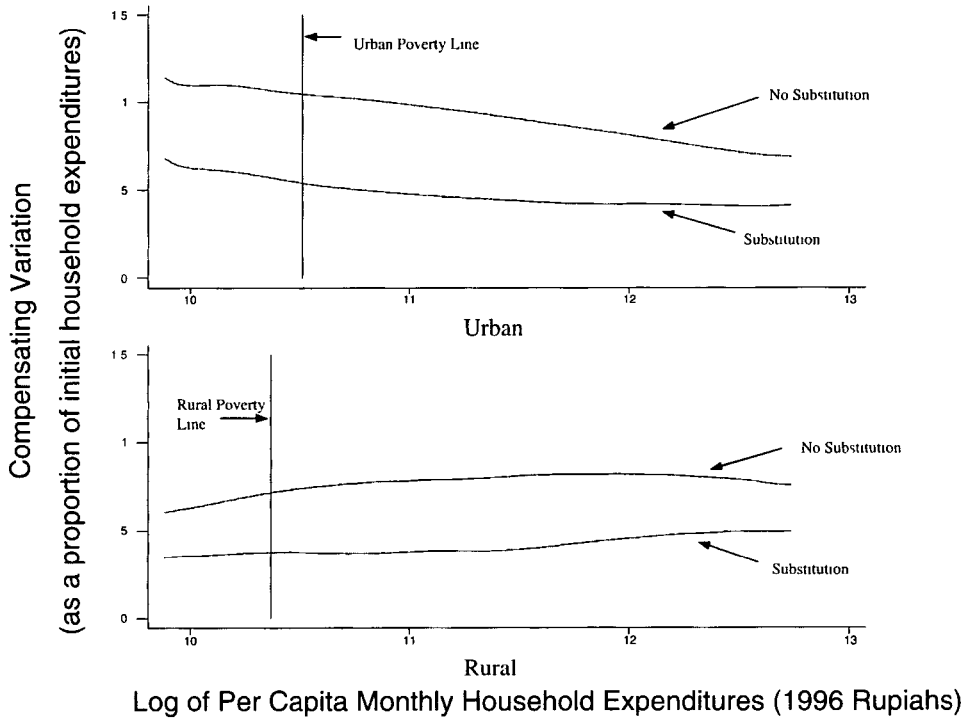
However, the differences in the levels estimated for expressions (2) and (5) are quite pronounced. The compensating variation measures that allow for substitution are substantially smaller than those that do not at all expenditure levels and for both urban and rural households. Indeed, as a rule of thumb, the estimates of expression (5) are roughly half the estimates of expression (2), with the difference greatest for lower-income urban households. Thus expression (5) suggests that the overall impacts of the crisis were not nearly as severe as suggested by expression (2).

Without further information it is difficult to know which of the results from expressions (2) and (5) are closer to the truth. We know that expression (2) surely overstates the impacts of the price changes, because it restricts households to consuming goods in the same proportions as they did before the large relative price changes of the crisis. However, we have reason to believe that expression (5) as currently estimated may dramatically understate the true compensating variation. If it does, the true compensating variation lies somewhere between the two regression lines for these expressions.

We believe the results for expression (5) may overstate the true degree of substitution because the reduction in food consumption implied by the ε matrix and the price changes results in very low caloric intakes—much lower than would actually be exhibited (and indeed has been suggested by measured changes in the body-mass index in Frankenberg, Thomas, and Beegle 1999). Essentially the problem lies with the estimated elasticities themselves. We believe these estimates may not be accurate for two important reasons. First and foremost is the fact that the estimated elasticities are essentially local approximations based on consumer behavior at the observed prices. Hence SUSENAS may give fairly good estimates of how households respond to a price change on the order of 5–10 percent. When the price changes are on the order of 100–300 percent, however, the answer is essentially dictated by choice of functional forms. This is troubling for

8. The functional forms used to estimate the elasticities are not flexible enough to allow the price responses to vary across the expenditure distribution. A more flexible approach may find that the poor substitute more than the wealthy because they are forced to economize on every price change. Alternatively, the wealthy may substitute more because their greater distance from subsistence gives them more opportunity to do so. If either of these possibilities is true, allowing the elasticities to vary by household expenditure level may yield different a distributional impact than that found with expression (5).

FIGURE 2. Compensating Variation, with and without Substitution Effects



most any parametric approach to the estimation of demand elasticities. In essence, we are forced to make out-of-sample predictions for every household; the farther the real price changes are from the range of prices (or unit values) in SUSENAS, the more important the choice of functional form becomes. A related problem is that our compensating variation calculations are averaged over all households in a per capita expenditure group, but our elasticity estimates include only those households with nonzero purchases of the various good aggregates. Ideally, we want an average price response over all households; excluding nonpurchasers from the elasticity estimates will overstate the offsetting effects of price substitution.

We still present results with the cross-price elasticities because, in principle, they are an important refinement over expression (2). Noting the difficulties of accurately accounting for substitution behavior given only one cross-section of households and price changes of the magnitude found in Indonesia in 1998, we do not claim that the true postcrisis compensating variations are those estimated from expression (5). We do find it reassuring that the distributional consequences implied by expression (5) are the same as those implied by espression (2) and present results from both specifications. The combined results from both expressions (2) and (5) may be of greater use to policymakers than either expression alone.

IV. Extensions

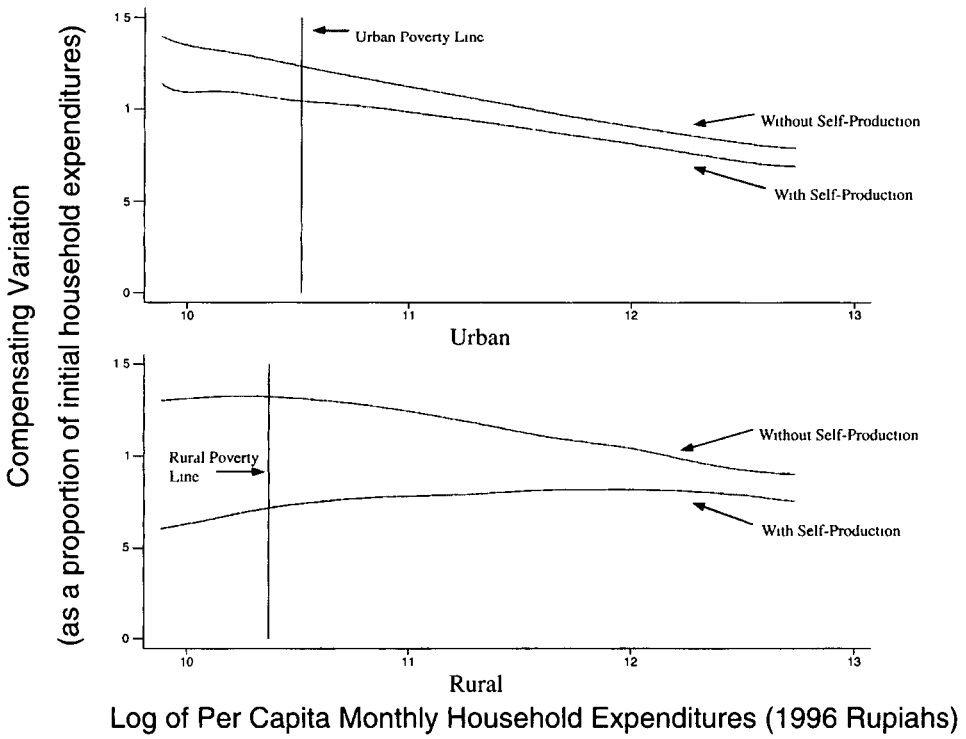
Having presented the basic results for expressions (2) and (5), we now turn to four extensions that explore the robustness of the findings. These extensions explore the effects of ignoring the services provided by owned housing and self-produced food, examine what happens if we use a smaller number of more highly aggregated goods instead of the 219 highly disaggregate goods, study the degree of spatial variation in the compensating variation measures, and look at how the compensating variation measures may be influenced by household size and demographic composition.

The first extension investigates differences in our findings if we do not account for the services provided by owned housing and self-produced food. Figure 3 presents this scenario, separately for urban and rural households, by depicting the nonparametric regression lines for the compensating variation given in expression (2) with and without valuing self-produced food and owned housing as negative expenditures. Ignoring household self-production dramatically changes the results, especially for rural households. For households in urban areas, where only a minority of households produce some of their own food, the qualitative results are the same whether or not we value self-production: poor urban households are affected substantially more than wealthy ones. However, without self-production and owned housing, the regression line is shifted upward in an almost parallel fashion, so that the levels of compensating variation are about 15 percent greater than before.

For rural areas, ignoring self-production results in attributing the greatest adverse consequences to the rural poor as opposed to the rural wealthy, a complete reversal of the findings in figure 1. The levels of compensating variation for the rural poor also increase dramatically, almost doubling to about 130 percent of initial expenditures from the 70 percent reported in table 3. The levels also rise for the rural wealthy but by a much smaller proportion. Clearly, the ability of rural households (especially lower-income rural households) to produce their own food buffered those households from the worst effects of the crisis. Urban households to a large degree could not share in this benefit.

We are also interested in exploring how the degree of aggregation or disaggregation affects the results. Remember that we attempted to match consumption and price changes at as low a level of aggregation as possible to allow more fully for heterogeneity in both consumption choices and price changes. The motivation, however, for looking at a more aggregate index stems from the fact that the disaggregated index accounts for only 79 percent of household expenditures on average. It is possible that we excluded important unmatched goods and that this exclusion can exacerbate or mitigate the measured welfare effects, depending on the relative price changes of those excluded goods. Concerned about this potential bias, we compute another compensating variation measure based on 19 aggregate commodities instead of the original 219. These aggregates include 15 food categories, such as cereals and meat, and 4 nonfood categories, such as housing

FIGURE 3. Compensating Variation, with and without Self-Production

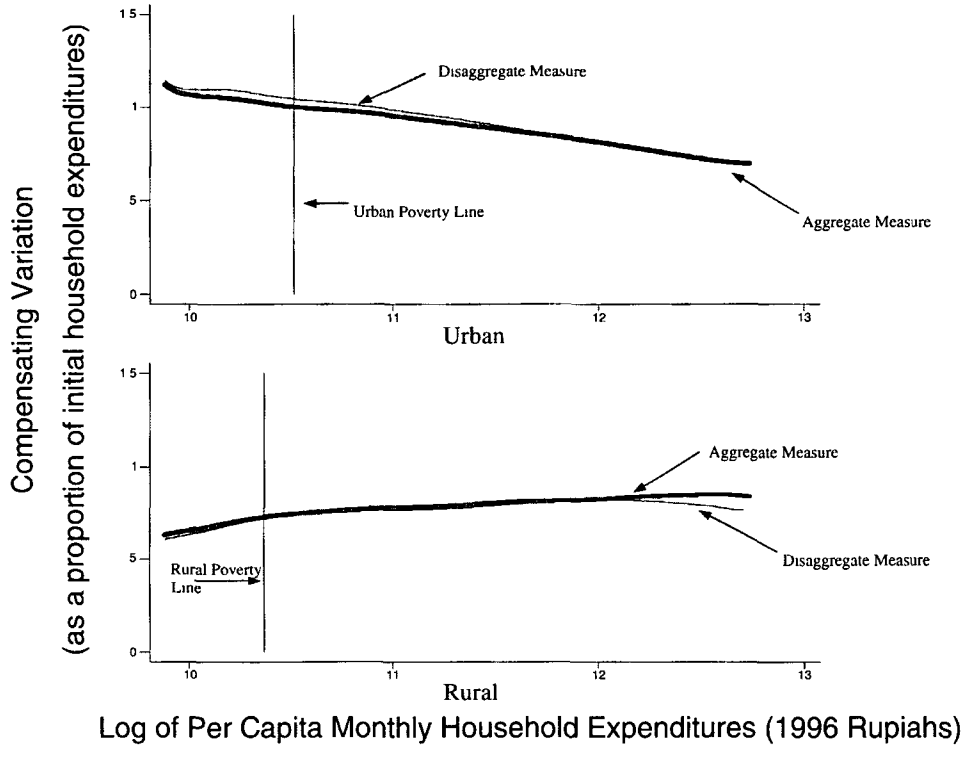


and clothing. A benefit of this aggregate measure is that it covers 97 percent of the average household's expenditures (this coverage is virtually the same for rural or urban households and across the income distribution).

The results suggest that little is changed if we base the compensating variation measures on the more aggregate consumption goods (figure 4). Indeed, the regression lines representing the aggregate and disaggregate measures are virtually identical for both urban and rural households. The analysis based on aggregated data is essentially unaffected by aggregation bias, at least in this case, where our disaggregate measures include many important consumption goods. We find this reassuring on two fronts. First, figure 4 implies that our main results are not biased by any "missing" consumption. Second, not every household survey records consumption at such a disaggregate level as SUSENAS. Figure 4 suggests that similar analysis conducted with other surveys may suffer little aggregation bias as long as the basic consumption categories are present in the data.

All of the preceding analysis has ignored cross-spatial variation in the compensating variation measures, except by distinguishing urban from rural households. However, Indonesia's population is spread out over 27 provinces on numerous islands. Many of the studies previously cited concerning postcrisis household changes have shown that different areas of the country were affected

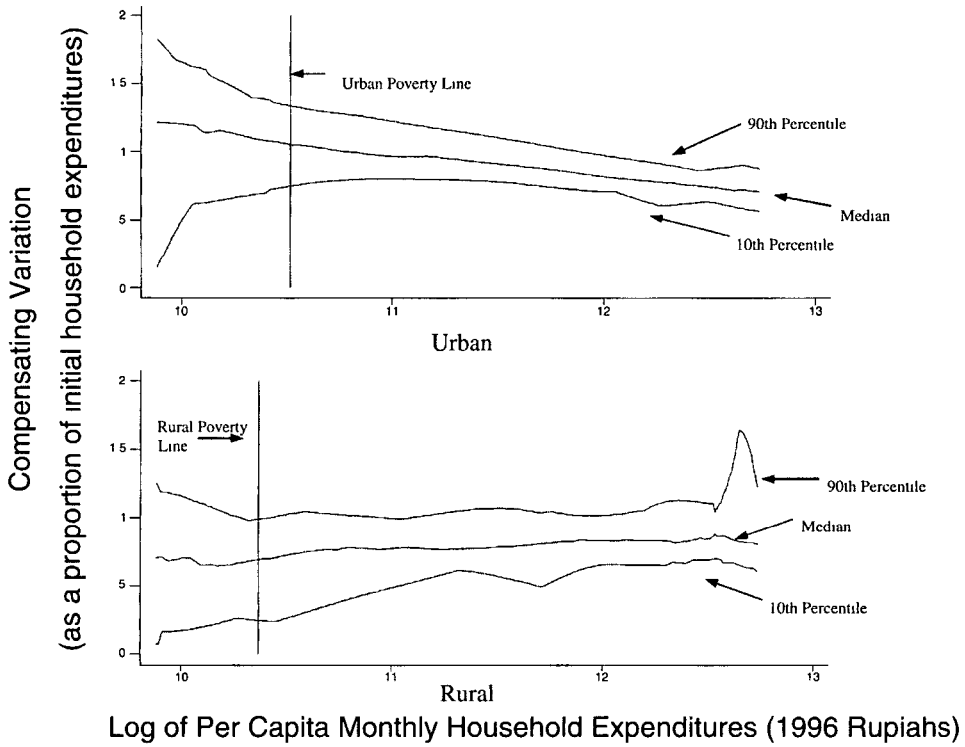
FIGURE 4. Compensating Variation, Aggregate and Disaggregate Measures



differently by the crisis due to geographic variation in both price changes and sources of income. Our findings are no different. When we calculate the mean province-level values of our main compensating variation measure, the geographical diversity is readily apparent. For example, households in urban East Nusa Tenggara, a collection of islands east of Bali, needed an additional 53 percent of pre-crisis expenditures to maintain consumption, whereas households in urban Southeast Sulawesi needed 124 percent. Although in every province rural households faced a smaller compensating variation than their urban counterparts, the regional variation among rural households is equally dramatic. In Bengkulu, a Sumatran province, the compensating variation for rural households averaged 105 percent, whereas the figure for Irian Jaya was only 30 percent.

To summarize the geographic results visually, we estimated the nonparametric regression lines separately for each province, ordered the estimated compensating variations at each point in the expenditure distribution, and plotted the 10th and 90th percentile of the province-specific compensating variations, along with the median. The resulting figure presents some measure of the geographic variance of the impacts while controlling for per capita household expenditures (figure 5). It is apparent that the effects of the crisis depended not only on the location of the household in the national expenditure distribution but also on

FIGURE 5. Dispersion of Compensating Variation across Provinces, 10th, 50th, and 90th Percentiles (with Self-Production)



the location of the household in space. For urban households the 90th percentile is roughly twice that of the 10th percentile, and this ratio is even greater for poorer households. Among rural households the spread between the 90th and 10th percentiles is even greater than that for urban households. Clearly, even within rural and urban areas, household location is an important determinant of the overall impact of the crisis.

Up to this point our principal measure of household welfare has been the household's per person expenditure level. Although a common measure, it imposes certain restrictions on how welfare may or may not vary across observable demographic information, such as household size or age and gender composition. Specifically, this measure does not recognize the possibility of scale economies at the household level or the fact that consumption needs of individual household members may vary across gender or the life cycle. Larger households, especially those with a larger number of working age adults, may be better off than smaller households at equivalent income levels because purchases of household public goods are shared among a greater number of household members. A consequence of this may be proportionally greater household expenditures for food (an important household private good), as public goods, such as housing, are more easily

afforded. In addition to household size, the demographic composition of the household is likely to affect household consumption choices to the extent that consumption needs vary across characteristics of household members. For example, households with children will almost surely spend more on education than otherwise equivalent households without children. Of course, differences in household consumption due to demographic influences will affect our compensating variation measures.

We explore these issues in our final extension with some simple ordinary least squares regressions of the main compensating variation measure on household size and demographic composition, as well as some relevant covariates, including per capita household expenditures (table 4). An earlier finding of this article is apparent in table 4 in the estimated coefficients for household expenditures: The positive coefficient for rural households indicates that the impact of the crisis increases with income levels in rural areas, whereas the opposite story is indicated by the urban household coefficient. Turning to the question of household size, larger households (especially in rural areas) are associated with higher compensating variations. The potential reasons for this result are numerous, but

TABLE 4. CV Regressions with Household Demographic Controls

Independent variables	Rural households		Urban households	
ln(Household PCE)	0.0919	0.0948	-0.1709	-0.1720
	0.0144	0.0147	0.0075	0.0076
ln(Household size)	0.1034	0.0722	0.0293	0.0013
	0.0105	0.0127	0.0067	0.0068
Proportion of household:				
male, 0–4 years old	—	0.1362	—	0.0817
		0.0401		0.0211
female, 0–4 years old	—	0.1224	—	0.0758
		0.0404		0.0235
male, 5–14 years old	—	0.0511	—	0.0289
		0.0300		0.0158
female, 5–14 years old	—	0.0186	—	-0.0106
		0.0307		0.0165
male, 15–59 years old	—	—	—	—
female, 15–59 years old	—	-0.0516	—	-0.0293
		0.0276		0.0139
male, 60 years or more	—	-0.0226	—	-0.0791
		0.0352		0.0303
female, 60 years or more	—	-0.1441	—	-0.1857
		0.1362		0.0276
R^2	0.1117	0.1131	0.2620	0.2677
Unweighted N	37493		24472	

Note: Ordinary least squares regressions include age, gender, and education of household head as well as provincial dummies. Standard errors, reported below the estimated coefficients, are corrected for observational dependence within survey clusters.

Source: Authors' calculations from 1996 SUSENAS and BPS price data.

surprisingly, higher food shares resulting from the larger household sizes is not one of these explanations. If anything, food shares are negatively related to household size (data not shown), especially in urban areas, once we control for per capita household expenditures. This finding may be somewhat surprising in light of the earlier discussion, but it is largely consistent with the multicountry results reported in Deaton and Paxson (1998). For whatever reasons, larger rural households tend to consume more of goods whose prices have disproportionately risen. The finding for urban households is the same, although not as pronounced. Indeed, once we control for household demographic composition the impact of household size on the compensating variation measure disappears for urban households.

The second columns in both the urban and rural panels of table 4 report the results from a regression of the compensating variation measure on the proportion of household members falling into eight age and gender categories: young (under age 5) boys and girls, children and adolescents (age 5–14), adult men and women, and male and female elders (age 60 and older). The excluded reference category is the proportion of adult men in the household. The results do indeed suggest that consumption patterns differ by age and, to a lesser extent, by the gender composition of the household members. Urban and particularly rural households with a large proportion of young children face a significantly higher compensating variation measure. Households with young children tend to spend more on food, especially rice; because the prices of these commodities rose the fastest, these households suffered disproportionately. Conversely, households with a higher proportion of adult women, especially elderly women (and elderly men in urban areas), tend to face lower compensating variations, in part reflecting the relatively low food needs of these groups. Thus in addition to urban/rural status, provincial location, and overall income, important factors that mediate the crisis impact at the household level include household size (in rural areas) and household composition.

V. Conclusions and Discussion

Analyzing the distributional impacts of economic crises is important and, unfortunately, an ever more pressing need. If policymakers are to intervene to help those most adversely affected, they need to identify those who have been hurt most and measure the magnitude of the harm they have suffered. Furthermore, policy responses to economic crises typically must be timely. We developed a simple methodology to fill the order and applied our methodology to analyze the impact of the Indonesian economic crisis on household welfare there. In particular, we estimated the compensating variation for Indonesian households following the 1997 Asian currency crisis. We found that virtually every household was severely affected, although the urban poor fared the worst. The ability of poor rural households to produce food mitigated the worst consequences of the high inflation. We found that the distributional consequences were the same

whether we allowed households to substitute toward relatively cheaper goods or not. Furthermore, these findings were not biased by any missing consumption. We obtained very different results, however, if we ignored the relative benefits of self-production or owned housing. Finally, even within urban or rural areas, the geographic location of the household mattered greatly, and households with young children suffered disproportionately adverse effects.

Although our methodology is simple and uses more or less readily available data, it is not perfect. Two limitations in particular need to be kept in mind. First, it is easy to forget that the economic crisis was not the only change affecting Indonesia's economy over this period. Concurrent with the crisis, some areas of Indonesia were hard hit by forest fires and others by drought. These disasters affected prices, so that not all the price changes we observe in the data are due solely to the economic crisis. Put another way, prices would have changed some even without the crisis. Our analysis speaks to the net effect of the many concurrent economic changes Indonesian households faced. We do not make any attempt to decompose what portion of the price changes were due to the financial crisis.

Second, all of the analysis concerns nominal changes. We focus on compensating variation because consumption is an important component of household welfare and we have detailed information on household consumption as well as detailed information on price changes for consumption goods. Given these initial conditions we believe we can look at compensating variation measures in a careful and nuanced manner. SUSENAS contains much less detail on sources of household income, and the available information on changes in factor incomes is also much less detailed. Our inability to accurately forecast changes on the income side renders us mute in terms of the real impacts of the crisis. However, in future work it may be possible to match our compensating variation measures with estimates of changes in nominal income in order to obtain a forecast of the "real" impacts of the crisis.[9] Multiple methods of forecasting or identifying vulnerability, combined with a greater understanding of the behavioral responses to crisis, will likely yield improved tools for policymakers hoping to understand and alleviate the effects of economic crisis.

How well our nominal measures predict actual outcomes is another topic of future research. For now we are able to compare our predictions with results from studies that have analyzed postcrisis household data. Several such studies present summary results broadly consistent with our predictions. Our method predicts greater consumption impacts in urban than in rural areas, as Frankenberg, Thomas, and Beegle (1999) have documented. Although the exact magnitude of the estimated expenditure decline depends on the particular price deflator adopted, Frankenberg, Thomas, and Beegle consistently find that urban house-

9 Robilliard, Bourguignon, and Robinson (2002) present a novel approach for estimating changes on the income side by matching a microsimulation model based on household data with a computable general equilibrium model.

holds suffered greater declines than rural ones. For example, when applying the BPS deflators (derived from the same underlying price change data used here), they find that real per capita expenditures declined 34 percent in urban households and 13 percent in rural households. Exploiting the panel nature of the data, the authors also find that larger households are more likely to enter poverty than are smaller households, a result suggested by our table 4.

The crisis impacts at the bottom of the expenditure distribution were also more severe for urban households than rural ones. Suryahadi and others (2000) found that between 1996 and 1999 urban poverty headcounts increased 152 percent (from 3.8 percent to 9.6 percent), whereas rural poverty increased only 57 percent (from 13.1 percent to 20.6 percent). An alternative poverty measure more sensitive to distributions among the poor increased 202 percent in urban areas and 84 percent in rural areas. Thus changes in the severity as well as the incidence of poverty were greater in urban areas.[10]

These studies also highlight behavioral responses that households undertook to cushion the effects of the crisis. Our compensating variation measures do not take into account these behavioral possibilities. Households were able to smooth consumption through two main mechanisms. First, households were able to adjust their labor supply in response to the changing conditions. These labor responses include women drawn into unpaid family work in household enterprises, perhaps the most common response observed by Frankenberg, Thomas, and Beegle (1999). Second, households with relatively liquid assets were also able to smooth consumption in relation to asset-poor households. Although the compensating variation measures are not currently flexible enough to incorporate these behavioral mechanisms, they nevertheless appear to retain merit as predictors of post-crisis outcomes.

If our measures have predictive value, they can aid in the design and targeting of policy responses. After 1997 Indonesia implemented several major policy programs to alleviate the adverse effects of the crisis. Our results would suggest that special attention needs to be paid to urban areas. However except for some employment programs, most policy responses did not target urban areas.[11] Furthermore, most of the programs contained no geographic targets for regions most affected by the crisis (Sumarto, Suryahadi, and Widyanti 2001). Because the dis-

10. The dynamic effects of the crisis are also important to consider when assessing changes in household welfare. The estimated poverty rates increased dramatically from August 1997 to August 1998, but by August 1999 they had fallen more than halfway to the 1997 levels. Because the record shows that prices rose rapidly and wages were slower to respond, our method appears most applicable at the start of a crisis, before factor returns can fully respond to the increase in commodity prices

11 The four most significant policy responses to the crisis were the following Rice subsidies were originally targeted to the poor but were subsequently shared more broadly; monthly scholarships at the primary and secondary level targeted the poor but were not widely available, block grants were provided to community health centers, with no concomitant rise in overall government spending on health; work programs from a variety of government organizations were initially located in urban areas, but made no special efforts to target the most poor.

tributional impacts of the crisis appear to vary spatially, our methods could also potentially aid the geographic targeting of policy responses.[12]

In sum, these measures of compensating variation appear to have some predictive power when applied to the 1997 Indonesian currency crisis. Equally important, they are relatively simple to estimate and are available as soon as data on commodity price changes are collected. Any attempt to comprehensively measure the real costs to households requires time and energy intensive data collection; the results of these efforts may be available long after policymakers have responded to the crisis with new or modified social policies. Because informational needs are immediate, the simple measures presented here should prove useful. Exactly how these measures predict actual outcomes remains a topic of ongoing research.

References

Benjamin, Dwayne. 1992. "Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models." *Econometrica* 60(2):287–322.

Bidani, Benu, and Martin Ravallion. 1993. "A Regional Profile of Poverty in Indonesia." *Bulletin of Indonesian Economic Statistics* 29(3):37–68.

Case, Anne C. 1991. "Spatial Patterns in Household Demand." *Econometrica* 59(4):953–66.

Chaudhuri, Shubham, and Martin Ravallion. 1994. "How Well Do Static Indicators Identify the Chronically Poor?" *Journal of Public Economics* 53(3):367–94.

Deaton, Angus. 1988. "Quality, Quantity, and Spatial Variation of Price." *American Economic Review* 78(3):418–31.

———. 1990. "Price Elasticities from Survey Data: Extensions and Indonesian Results." *Journal of Econometrics* 44(3):281–309.

———. 1997. *The Analysis of Household Surveys*. Baltimore, Md.: Johns Hopkins University Press

Deaton, Angus, and John Muellbauer. 1980. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.

Deaton, Angus, and Christina Paxson. 1998. "Economies of Scale, Household Size, and the Demand for Food." *Journal of Political Economy* 106(5):897–930.

Fan, Jianqing. 1992. "Design-Adaptive Nonparametric Regression." *Journal of the American Statistical Association* 87(4):998–1004.

Frankenberg, Elizabeth, Duncan Thomas, and Kathleen Beegle. 1999. "The Real Costs of Indonesia's Economic Crisis: Preliminary Findings from the Indonesia Family Life Surveys." RAND Working Paper No. 99-04. Santa Monica, Calif.

12. Although poverty changes have been more severe in urban areas, overall poverty levels are greater in rural areas. This, of course, begs the question of whether spending and relief programs should be targeted to the "poor" or to the "shocked " In that sense our methods are suited to the design of safety "ropes" as opposed to safety "nets" insofar as these methods identify households most affected by change, not necessarily those with the lowest overall welfare (see Sumarto, Suryahadi, and Pritchett 2000).

Friedman, Jed. Forthcoming. "How Responsive Is Poverty to Growth? A Regional Analysis of Poverty, Inequality, and Growth in Indonesia, 1984–1999." In Ravi Kanbur and Anthony Venables, eds., *Spatial Inequality and Development*. Oxford: Oxford University Press.

Levinsohn, James, Steven Berry, and Jed Friedman. Forthcoming. "Impacts of the Indonesian Economic Crisis: Price Changes and the Poor." In Michael Dooley and Jeffrey Frankel, eds., *Managing Currency Crises in Emerging Markets*. Chicago: University of Chicago Press.

Ravallion, Martin. 1994. *Poverty Comparisons. Fundamentals in Pure and Applied Economics*. Vol. 56. Chur, Switzerland: Harwood Academic Press

Ravallion, Martin, and Benu Bidani. 1994. "How Robust Is a Poverty Profile?" *World Bank Economic Review* 8(1):75–102.

Ravallion, Martin, and Dominique van de Walle. 1991. "The Impact on Poverty of Food Pricing Reforms: A Welfare Analysis for Indonesia." *Journal of Policy Modeling* 13(2):281–99.

Robilliard, Anne-Sophie, François Bourguignon, and Sherman Robinson. 2002. "Crisis and Income Distribution: A Micro-Macro Model for Indonesia." World Bank. Washington, D.C.

Smith, James, Duncan Thomas, Elizabeth Frankenberg, Kathleen Beegle, and Graciela Teruel. 2002. "Wages, Employment, and Economic Shocks: Evidence from Indonesia." *Journal of Population Economics* 15(1):161–93.

Sumarto, Sudarno, Asep Suryahadi, and Lant Pritchett. 2000. "Safety Nets and Safety Ropes: Who Benefited from Two Indonesian Crisis Programs—the 'Poor' or the 'Shocked'?" World Bank Policy Research Working Paper No. 2436. Washington, D.C.

Sumarto, Sudarno, Asep Suryahadi, and Wenefrida Widyanti. 2001. "Designs and Implementation of the Indonesian Social Safety Net Programs: Evidence from the JPS Module in the 1999 SUSENAS." SMERU Working Paper. Jakarta.

Suryahadi, Asep, Sudarno Sumarto, Yusuf Suharso, and Lant Pritchett. 2000. "The Evolution of Poverty During the Crisis in Indonesia, 1996–1999." Policy Research Working Paper No. 2435. World Bank, Washington, D.C.

Thomas, Duncan, Kathleen Beegle, Elizabeth Frankenberg, Bondan Sikoki, John Strauss, and Graciela Teruel. 2001. "Education in a Crisis." University of California–Los Angeles, Department of Economics.

WHO (World Health Organization)/FAO (Food and Agriculture Organization of the United Nations)/UNU (United Nations University) Expert Panel. 1985. *Energy and Protein Requirements*. WHO Technical Report No. 724. Geneva.

THE WORLD BANK ECONOMIC REVIEW, VOL 16, NO. 3 425–448

Density versus Quality in Health Care Provision: Using Household Data to Make Budgetary Choices in Ethiopia

Paul Collier, Stefan Dercon, and John Mackinnon

Usage of health facilities in Ethiopia is among the lowest in the world; raising usage rates is probably critical for improving health outcomes. The government has diagnosed the principal problem as the lack of primary health facilities and is devoting a large share of the health budget to building more facilities. But household data suggest that usage of health facilities is sensitive not just to the distance to the nearest facility but also to the quality of health care provided. If the quality of weak facilities were raised to that currently provided by the majority of facilities in Ethiopia, usage would rise significantly. National data suggest that given the current density and quality of service provision, additional expenditure on improving the quality of service delivery will be more cost-effective than increasing the density of service provision. The budget allocation rule presented in the article can help local policymakers make decisions about how to allocate funds between improving the quality of care and decreasing the distance to the nearest health care facility.

This article combines household survey data on health care choices in rural Ethiopia with budget data on the costs of health provision to analyze the tradeoff between the density of service provision and its quality. We develop a budget allocation rule that can be used in local decisionmaking. National data suggest that given the current density and quality of service provision, additional expenditure on improving the quality of service delivery will usually be more cost-effective than increasing the density of service provision.

The focus is on the usage of health services rather than health outcomes.[1] The link from usage to outcomes is complex. For example, Ethiopian health care

Paul Collier is Director of the Development Economics Research Group at the World Bank His e-mail address is pcollier@worldbank.org. Stefan Dercon is a University Lecturer in Economics at Oxford University, linked to the Centre for the Study of African Economies. His e-mail address is stefan.dercon@economies.ox.ac.uk. John Mackinnon is an independent researcher and advisor on development issues. His e-mail address is jwm1962@aol.com.

1. Health outcome data, in particular adult nutrition and functioning data, are also available but were not included in the analysis. In principle, the links among distance, quality of health services, and outcomes can be fruitfully analyzed (examples are in Lavy and others 1996, and Thomas, Lavy, and Strauss 1996). However, strong seasonality in nutritional outcomes and the importance of controlling for health

DOI: 10.1093/wber/lhf003

providers use the opportunity provided by curative visits to provide preventative treatment, so that the impact of health care on the illness that motivates the visit is likely to underestimate the overall contribution of usage to health.

In Ethiopia usage is much lower than in most other low-income countries. In 1996 only 9 percent of the population reported visiting a modern health facility during the past two months (Dercon 2000). This compares with 14 percent in Kenya, also based on a two-month recall period (Appleton 1998); 13 percent in Ghana, based on a four-week recall period; and 15 percent in Côte d'Ivoire, based on a four-week recall period (Lavy and Germain 1994, Dor and Van der Gaag 1988). These differences exist despite far worse health outcomes in Ethiopia.[2] Among those who sought treatment, the vast majority relied on public facilities—not surprising in a country in which in 1999 more than 90 percent of rural health clinics (centers and stations) were owned by the government (Ministry of Health 1999). This reliance on public facilities distinguishes Ethiopia from other Sub-Saharan countries. In Côte d'Ivoire, for example, only about 40 percent of households relied on public facilities; in Kenya only about a third did so. Thus there is a reasonable presumption that poor health outcomes are partly caused by the atypically low usage of health facilities and that policy toward public provision will continue to be important in rectifying it.[3]

Policy to increase usage in Ethiopia currently focuses on increasing the physical coverage of facilities. This approach is based on the belief that distance is the main disincentive to using services. Increased coverage is very costly in terms of both the capital costs of building facilities and the fixed recurrent costs of running them. The policy thus squeezes variable recurrent expenditure per facility, reducing staffing and the supply of material inputs, such as drugs, and worsening the quality of service provision. In a highly revenue-constrained environment, there is a sharp tradeoff between reducing the distance to health facilities and improving their quality.

To the extent that existing facilities have spare capacity, increasing the quality of services may increase usage. A dollar spent on increasing the quality of service provision may increase usage by more than a dollar spent on increasing the number of facilities. In this case, aside from considerations of equity, improving quality dominates building new facilities because there is a larger increase in usage and both existing and new users get better services. Only if the reverse holds (so that services for new users are at the expense of services for existing users), is there a quantity–quality tradeoff.

---

endowments make this analysis beyond the scope of this article. For a detailed analysis of the dynamics of adult nutrition in Ethiopia and the link with illness and health care, see Dercon and Krishnan (2000).

2. The *World Health Report* (WHO 2000) estimates male under-age-five mortality rates at 100 per thousand for Kenya, 145 for Côte d'Ivoire, and 188 for Ethiopia, with adult mortality rates also higher in Ethiopia. In 1994 life expectancy in Ethiopia was estimated to be only about 43 years for men—well below the African average of 52 years.

3. Nevertheless, in the econometric model, the response to increased spending on public services and the impact on usage of private health providers will have to be taken into account (Filmer, Hammer, and Pritchett 1997).

The responsiveness of health care demand to the distance from facilities is well established. By contrast, until recently, demand responsiveness to the quality of health care has been largely ignored. Recent work has pointed to the importance of including quality variables. (See, for example, Akin and others 1986, Hotchkiss 1993, Lavy and Germain 1994, Lavy and others 1996, Litvack and Bodart 1993, 1995, Mwabu, Ainsworth, and Nyamete 1993, Thomas, Lavy, and Strauss 1996. For an overview see Alderman and Lavy 1996.) Distance and quality are now recognized to be jointly important in determining usage.

This article applies similar techniques in Ethiopia. The distance–quality tradeoff has not previously been studied in such a deprived environment. The article further innovates by relating household behavior to budgetary expenses. To date the main policy focus of research on the demand effects of health care quality has been on user fees: if fees raise quality sufficiently, they can enhance usage. This article is concerned with the budgetary choice between government expenditures on capital programs versus recurrent health expenditures. That is, it takes total health care financing as given. Clearly, in the distance–quality tradeoff the main choice is one of budget allocation rather than cost recovery. By focusing on the distance–quality tradeoff we do not mean to imply that this is necessarily the most important issue in the allocation of public expenditure on health care. However, it has been neglected, because it is less obvious than tradeoffs such as clinics versus hospitals.

Section I formalizes the determination of demand by density and quality and discusses the relation between demand and welfare. Section II describes health provision in Ethiopia. Section III presents regression results on the demand for health care, comparing the effects of the quantity and quality of the provision of clinics on the quantity of health service usage. Section IV estimates budgetary costs of increased quantity of provision and improved quality of services and derives the relative cost-effectiveness of capital and recurrent expenditures. Section V summarizes the article's main findings.

I. A Formal Framework for Analyzing the Effects of Quality and Distance on Usage of Health Care

The government is assumed to want to maximize health outcomes subject to a budget constraint and the behavior of the representative household. The instruments the government has are the quality and the density of the primary health care network. This is a principal-agent problem in which the government needs to take into account the effects of its instruments on household demand for health care. Households are assumed to be characterized by a utility function in which health is an argument. This is the simplest way to represent the value placed by households on health in a static framework:[4]

4. The alternative—incorporating the likely effects of health on future income—would require an explicitly dynamic model. For an extensive survey on modeling the demand for health care, see Behrman and Deolalikar (1990) and Strauss and Thomas (1995).

$$U = U(H,C) \tag{1}$$

where H is a measure of health and C is consumption excluding health. Time may be spent either in providing labor or in visiting health facilities:

$$L = l + t(D).V \tag{2}$$

where L is the endowment of time, l is labor, D is the distance from facilities, t is the time spent per visit to a clinic, and V is the household's demand for visits to health facilities. We could simply assume that time is a linear function of distance; however, doing so ignores the possibility that people substitute into faster forms of transport as distance increases. In this article we use data on distance rather than time, and the form of the function is not important. The full income budget constraint is given by

$$w.L = C + F.V + w.t(D).V \tag{3}$$

where w is the real wage rate and F is the user charge for health care.

The amount of health actually achieved is determined by a production function in which both demand and the quality of facilities play a part:

$$H = H(C,V,Q) \tag{4}$$

where Q is quality.

For simplicity these functions are treated as deterministic; the stochastic nature of health outcomes is important in many contexts but does not affect this model.

The household chooses L and V to maximize equation (1) subject to equations (2–4). This problem yields demand functions both for health and for health care:

$$H = h(D,Q,F,w) \tag{5}$$

$$V = V(D,Q,F,w). \tag{6}$$

Our concern is with the demand functions, in particular the partial derivatives of the demand function for health care with respect to distance and quality. Empirical work has shed light on other derivatives of the demand function. Gertler and van der Gaag (1990) use the effects of distance on usage to estimate the effects of price on usage; Mwabu, Ainsworth, and Nyamete (1993) use the effects of seasonal movements in the cost of time to isolate the substitution effects of increases in the cost of time, finding that people visit clinics less in the busy season. Quality effects have been researched by Lavy and Germain (1994), Lavy and others (1996), and Thomas, Lavy, and Strauss (1996). Although changes in usage do not directly tell us about the change in welfare, they do have implications about the effectiveness of medical care. Because usage of health care has an opportunity cost in terms of consumption, we can distinguish between an income effect and a pure substitution effect in the choice between health outcomes and consumption. A reduction in distance

is a price reduction in usage and, as a result, in health outcomes. As long as health is a normal good, both the income and substitution effects work to increase the consumption of health outcomes and hence of usage. By contrast, an improvement in quality does not reduce the cost of usage but rather its effectiveness. Hence fewer visits will be needed to maintain a given health outcome. Improved quality thus directly lowers the price of health outcomes. Although there will be both an income and a substitution effect increasing the consumption of health outcomes, this need not imply increased usage of health services. Increased usage of health services would require reduced consumption of other goods and so will come about only if the substitution effect into improved health is sufficiently strong to outweigh the income effect, which tends to increase consumption of other goods and so reduce visits.

This can be seen in a simplified example of the model. Let the health production function (equation [4]) take the form

$$H = V.Q \tag{7}$$

so that we abstract from any effect of consumption onto health. Let the opportunity cost of visits in terms of consumption take the form

$$C = w.L - D.V \tag{8}$$

so that we abstract from user charges and make the cost of distance linear in distance. Finally, let the utility function take the form

$$U = H.C. \tag{9}$$

The optimal number of health care visits is now:

$$V^* = w.L / 2D. \tag{10}$$

The demand for visits is thus invariant with respect to quality while being a decreasing function of distance. In this case the income and substitution effects of a change in quality precisely net out. This is not a general result but depends on the functional form of the health production function and the utility function. For example, if the health production function takes the form

$$H = V^Q \tag{11}$$

then the optimal number of health care visits becomes

$$V^* = \frac{Q.w.L}{(1+Q)D} \tag{12}$$

so that the derivative with respect to quality becomes strictly positive, whereas the derivative with respect to distance remains strictly decreasing. (Note that the quality index developed later is bounded by 0 and 1, so that equation [11] is concave in visits.)

Given the objective of maximizing health outcomes, the government's allocation of spending between recurrent and capital spending must be such as to gen-

erate a level of quality and a density of provision that maximize health outcomes for given overall expenditures.[5] Formally,

$$G = r.Q - c.D \tag{13}$$

where G is the government health budget, r is the unit price of an improvement in quality, and c is the unit price of a reduction in the distance to the health facility.

The government maximizes H subject to its budget constraint (equation [13]) and the demand functions for health and health care (equations [5] and [6]). It must equalize the marginal effects of its expenditures on quality and density on the health outcome of the representative household. That is,

$$\frac{\frac{dH}{dQ}}{\frac{dH}{dD}} = -\frac{r}{c} \tag{14}$$

In general, expenditure on quality affects health outcomes through two routes, a direct increase in the efficacy of a given level of usage and an indirect change in the amount of usage, which as we have seen can be positive or negative. That is,

$$\frac{\frac{dH}{dQ}}{\frac{dH}{dD}} = \frac{\frac{\delta H}{\delta V}\cdot\frac{dV}{dQ} + \frac{\delta H}{\delta Q}}{\frac{\delta H}{\delta V}\cdot\frac{dV}{dD}} \tag{15}$$

$$= \frac{\frac{dV}{dQ}}{\frac{dV}{dD}} + \frac{\frac{\delta H}{\delta Q}}{\frac{\delta H}{\delta V}\cdot\frac{dV}{dD}} \tag{16}$$

Because the second term of expression (16) is negative, the optimum must be characterized by

$$-\frac{r\frac{dV}{dQ}}{c.\frac{dV}{dD}} < 1 \tag{17}$$

In the empirical application in section III we are not observing health outcomes but the differential impact of quality and density on usage. We can nevertheless test the efficiency of allocation of health care expenditures by testing whether they meet the condition set out in expression (17). That is, because improving quality raises health outcomes both through an efficacy effect and a usage effect whereas reducing distance works only through a usage effect, effi-

5. The implications of the government maximizing utility rather than health outcomes are more complicated to analyze.

cient allocation will require that the marginal impact on usage of a unit of expenditure on improved quality should be less than the marginal impact of a unit of expenditure on increased density.

Finally, we introduce a complication that reflects the dual nature of health care provision in Ethiopia between the public and private sectors. Both sectors maintain rural facilities, even if 90 percent of the modern facilities are public. The government can directly improve quality and reduce distance only with respect to its own facilities. The government decision problem remains as characterized, conditional on the demand for health and health care by private agents, though with V redefined as total usage. However, there is an additional substitution effect from the usage of private facilities. How such a substitution is valued by the government depends on the efficacy of private provision. If private facilities are of lower quality than public facilities, condition (17) is unaffected. If private facilities are of better quality than public facilities, an increase in overall usage would be the net effect of a reduction in the usage of good facilities more than offset by an increase in the usage of poor facilities. Such a change in usage could even reduce overall health states, so that dH/dV could potentially be negative. This would reverse the inequality in condition (17) as a condition for allocative efficiency but reinforce its core result. Because the effect of increased overall usage, V, on health outcomes would now be negative, the government would need to reduce its expenditure on distance reduction. Indeed, expenditure on distance reduction should be reduced until dH/dV becomes positive, at which point condition (17) would again apply. Expenditure on distance reduction greater than that implied by condition (17) is always excessive.

II. Health Care Provision in Ethiopia

Between 1987 and 1996 annual per capita government spending on health facilities averaged just $1.40 in constant 1987 U.S. dollars—far below the African average and indeed among the lowest in the world (World Bank 1996). Given these low levels of expenditure it is not surprising that provision is inadequate in terms of both density and quality. Facilities are thinly and unevenly spread, so that many households are far from clinics. Recent national sample data suggest that about 40 percent of rural households live more than 10 km from the nearest health facility (Central Statistical Authority 1999).

The services provided are often deficient. Clinics lack drugs and basic equipment. A government survey of facilities found that most lacked more than a quarter of the drugs rated as essential and a quarter lacked a refrigerator in which to store them.[6] These indicators appear worse than in other African countries.

6. The data—the most recent official statistics available—are from the sector investment office of the Ethiopian government and are based on a sample survey of 57 government health stations and centers in rural areas in 1995.

In Nigeria only about 28 percent of public health facilities lack drugs (Akin and others 1995); in Ghana about a third of public health facilities lack drugs (Lavy and Germain 1994). In Kenya public facilities lack antibiotics on average for about a tenth of the year (Mwabu, Ainsworth, and Nyamete 1993).[7]

The budgetary choice between the quantity of facilities and the quality of service provided is complicated in Ethiopia by two institutional features. First, the quantity of facilities, determined by the capital budget, is controlled by the Ministry of Planning and Economic Development, whereas the quality of services, determined by the recurrent budget, is controlled by the Ministry of Finance. The choice is liable to be more severely politicized than it would be if a single ministry were responsible for both types of expenditures. Second, most donors are precluded from funding recurrent expenditure and so, except through fungibility, increased donor funding is liable to shift the balance of expenditures toward facility provision rather than service quality. The tradeoff is intensified because capital expenditure not only has an opportunity cost in terms of contemporaneous recurrent expenditure but alters the subsequent composition of recurrent expenditure. Facilities have to be staffed, putting pressure on nonsalary recurrent expenditures.

The empirical analysis is based on a rural household survey conducted between 1994 and 1997 (see Dercon and Krishnan 1998 for details). Data are available from several rounds covering a panel of about 1,450 households and about 9,000 individuals. Detailed information on quality is available only for the first survey period (1994/95), except for a more limited number of villages (and therefore smaller sample), in which the health facilities were resurveyed just before the 1997 data collection round. We use mainly the data for the earlier period, but then use the smaller sample from 1997 to run fixed effects regressions, with varying quality, providing us with estimates for the quality effects, controlling for unobserved factors, such as fixed program placement effects.

The sampling frame was based on the major agro-ecological zones, focusing on the cereal-plough zones in northern, central, and southern Ethiopia; the *enset* complex; and the grain-hoe complex. Pastoralists and smaller zones were not included in this study. The sampled households cover 15 villages in an attempt to capture the different circumstances within different parts of each of the zones. The actual number of households chosen in each village reflects the population proportions. The relatively small number of villages was inspired by the practical constraints and costs of constructing high-quality longitudinal data and the desire to allow more detailed collection of community-level data. Attrition rates between 1994 and 1997 were about 5 percent. The sites chosen are very diverse in terms of wealth, agricultural resources, infrastructure, service provision, and so on. The resulting sample can be considered broadly representative of the population in these agro-ecological zones, covering more than 80 percent of the

7 Alderman and Lavy (1996) and Filmer, Hammer, and Pritchett (1997) discuss the problems with defining appropriate quality measures. Measures of infrastructure, drugs, and staff are commonly used.

rural population, although care has to be taken in interpreting the results. The sample is more suited for understanding causal relations across different types of households and the evolution over time of indicators than the actual levels of, say, the regional distribution of disease or malnutrition.

The survey collected comprehensive data on health and a large number of individual, household, and community characteristics. The health data were self-reported and covered incidence and duration of illness, symptoms, and actions taken. The recall period was the preceding four weeks. On average about 10 percent of the population reported an illness episode and 52 percent sought treatment (table 1).[8]

Richer households are only slightly more likely to consult someone for treatment (50 percent for the richest quartile versus 44 percent for the poorest group). The poor rely disproportionately on pharmacies and the nearest clinics, whereas richer households rely on hospitals and more distant facilities.[9]

Hospitals, pharmacies, and clinics in Ethiopia can be owned by private agents, the government, or NGOs. In the survey, households were asked to specify whether the facility visited was a government facility, a private (and modern) or NGO facility, or a traditional site. In many areas private hospitals or clinics may be at a considerable distance, so there may be no effective choice as to whether public or private facilities are used. The state sector still strongly dominates health provision. About 75 percent of hospitals and 90 percent of health stations and centers belong to the government. More than 90 percent of pharmacies are private, however. In our sample the nearest health centers and clinics are almost always public. Hence the choice as to public or private facility is largely subsumed in the choice between pharmacies and other facilities.

Table 2 shows the distribution of visits to private and public facilities.[10] More than 60 percent of the visits are to government facilities.[11] The apparent lack of difference across consumption quartiles conceals the fact that for the poorest

8. As in other surveys, illness episodes are biased as a result of self-reporting. For example, there was very little difference between rich and poor households in the incidence of reported illness. Educated parents reported more illnesses of their children. The results appear consistent with evidence from the Welfare Monitoring Survey collected in 1995–96 by the Central Statistical Authority. That survey is a nationally representative survey covering 7,000 rural and 4,000 urban households in 894 enumeration areas. It found that 19.4 percent of respondents in rural areas reported having been ill during the two months before the survey and that about 49 percent of them reported having sought treatment

9. The Ministry of Health and many studies focusing on health facility usage consider visits to pharmacies not as visits to "modern facilities" but rather as forms of self-treatment, because in principle, diagnosis and treatment should be restricted to health clinics, facilities, or trained health workers In the survey, households clearly did not make this distinction. In the econometric analysis, the implication of using a treatment definition including or excluding pharmacies is explored.

10. Data exclude traditional treatment. Private facilities include NGOs, which account for about 15 percent of visits to private facilities in our sample.

11. Data from the Welfare Monitoring System appear to suggest an even larger share of the private sector, with only 41 2 percent of those seeking treatment visiting a public facility in rural areas.

TABLE 1. Venue of Treatment Sought When Ill, by Consumption Quartile (percent)

Item	Poorest	Lower middle	Higher middle	Richest	All
No treatment	54.6	52.0	51.7	49.8	52.0
Nearest clinic	21.6	18.9	20.4	19.7	20.1
Nearest hospital	4.7	8.6	6.5	9.5	7.4
Hospital/clinic (not nearest)	4.8	6.1	6.3	6.7	6.0
Home of health worker	5.0	4.1	5.0	6.4	5.1
Pharmacy	8.4	9.2	8.5	6.1	8.0
Traditional healer	1.0	1.2	1.7	1.9	1.4

Source: Ethiopian Rural Household Survey, rounds 1–4.

group, pharmacies (the lowest tier of primary health care) are by far the most important source of private treatment, whereas for the richer groups, private health care includes a high proportion of visits by health workers at home or visits to private health clinics and centers.

III. ESTIMATION RESULTS

We use (reduced-form) demand relations for health care usage, distinguishing between distance to the facility, household and individual characteristics, and the quality of care, as in equation (6). People in different communities face differences in location and quality, even though the prices charged by clinics are uniform, some services, such as immunization, being free, and some having a low price (such as a 0.50 birr registration fee). Distances to health facilities are used to proxy the opportunity costs of getting to these services. Health stations or centers, the most important primary health care focal point in the health care system, are on average about 7 km from the villages; the nearest hospital is about 40 km.[12]

We include as explanatory variables a set of individual and household characteristics. Household income is proxied by household expenditure. We also control for the age and sex of the individual, as well as the age and sex of the household head and whether the mother completed primary school. Mean household expenditure in the sample is close to about $120 per capita per year (63 birr per capita per month in 1994 prices). The education variable shows the extremely low levels of schooling in rural Ethiopia, with only 2 percent of mothers having completed primary schooling.

12. The mean distance to the nearest center—about 7 to 8 km—is quite close to the average distances reported in the Welfare Monitoring System data in a nationally representative rural sample. The sample largely excludes pastoral and other lowland areas, which would have increased distance substantially. This mean distance is similar to the distance to public facilities in Kenya (8 km in the data used in Mwabu, Ainsworth, and Nyamete 1993) and in Ghana (8.3 km in the data used by Lavy and Germain 1994).

TABLE 2. Private and Public Facilities Visited for Treatment, by Consumption Quartile
(percent)

Type of facility	Poorest	Lower middle	Higher middle	Richest	All
Public	62.7	63.2	57.2	63.7	61.8
Private	37.3	36.8	42.8	36.3	38.2

Source: Ethiopian Rural Household Survey, rounds 1–4

In addition to these standard explanatory variables, we introduce information related to the quality of health care provided in the health facilities available to the households in the sample. Measuring quality in health facilities is always difficult (Alderman and Lavy 1996, Filmer, Hammer, and Pritchett 1997, and Thomas, Lavy, and Strauss 1996). As in most other contributions to this literature, we focus on attributes that are necessary for the quality of care to be adequate rather than on quality differences above such a threshold. To provide adequate quality of service, a clinic requires functioning equipment, qualified staff, and a reliable supply of material inputs, such as dressings and drugs. We proxy each of these three attributes. The proxy for functioning equipment is whether the clinic has a functioning refrigerator with a back-up power supply. Two-thirds of public clinics had such a refrigerator. The proxy for the quality of staff is whether the clinic has a qualified nurse in regular attendance. In 1994 about two-thirds of the sample of government facilities had a nurse in regular attendance. The proxy for material inputs is whether the clinic reports receiving a regular supply of antibiotics. The current system for distributing drugs to primary facilities is indeed likely to produce large quality differences. Government- and donor-financed drugs pass through distinct and uncoordinated systems, neither of which is responsive to needs. Primary (public) facilities do not purchase drugs from distribution systems, but rather receive such drugs as they are assigned. The lack of responsiveness of drug supply to need results in both persistent shortages and the persistent supply of drugs to locations where they are useless. Hence unsatisfactory quality is most likely the result of random failures in planned allocations (see Government of Ethiopia 1997). Only half the public clinics in the sample had a regular supply of antibiotics in 1994.[13]

There is wide variety in quality across facilities (table 3). Private health centers (including those run by NGOs) generally have better quality, but there are few of them, so that in many villages the nearest (and relevant) private facility is

13. Although the sample of health facilities is small, these measures are very close to the estimates from the 1995 survey by the sectoral investment office of the government of Ethiopia, which found that a quarter of public health stations and centers had no refrigerator and about half the public primary facilities lacked more than a quarter of essential drugs.

TABLE 3. Frequency Distribution of Public and Private Facilities, by Level of Quality
(percent)

	1994		Panel 1994–97			
Item	Public	Private	Public 94	Public 97	Private 94	Private 97
Equipment, inputs, and staff	31	50	0	21	50	50
Equipment and inputs only	0	0	0	0	0	0
Equipment and staff only	5	0	0	0	0	0
Inputs and staff only	16	0	36	36	0	0
Equipment only	16	3	21	29	0	0
Inputs only	12	0	29	0	0	0
Staff only	11	0	0	0	0	0
None	10	47	14	14	50	50

Note: Panel data for 1994–97 are based on four villages. In villages in which the only source of primary health care was a pharmacy (drug vendor), data refer to pharmacies. None of the pharmacies had a regular supply of antibiotics, a working refrigerator, or a regularly present, qualified staff (health worker, nurse, or doctor). Private facilities in the panel villages experienced no change in the quality indicators, so only one column is reported.

Source: 1994 and 1997 community-level surveys linked to the Ethiopian Rural Household survey.

a private drug vendor with very low quality. Only about a third of government facilities satisfy all the minimum quality criteria investigated.

Data on both facility characteristics and household characteristics were collected in 1994 from 15 villages. The reference period for facilities was for what was "typical" over the preceding 12 months. The households were surveyed three times during this reference period. The pooled dataset for these three rounds was then used in the regressions (allowing for the intertemporal dependence of errors using a random effects model). In 1997 both the facility and the household surveys were repeated in four villages. Because we have data on changes in quality for this subsample, we are able to test the robustness of the apparent effects of quality found in the cross-section analysis. In particular, by running fixed effects regressions, we can control for unobserved individual and community characteristics that may be correlated with health care quality (and other variables in the regression). If fixed program placement effects (such as the quality of health care being better in intrinsically wealthier areas populated with individuals with better health endowments) are present, the coefficients on quality variables in cross-section data are biased (see Rosenzweig and Wolpin 1986 for a classic discussion). Fixed effects regressions control for these fixed program placement effects, solving this problem. Because the coefficients are crucial for the analysis, the fact that we can do this is important, even though the subsample is relatively small. Although the potential for bias is well understood, such robustness tests are very rare because of lack of data. Note that neither the quality of private facilities nor the distance to health facilities changed in the subsample between 1994 and 1997 (no new building took place), so that the robustness of

the results can be checked only with respect to the effect of the quality of public facilities.

Several different models are estimated to explain usage of modern health care services when one is ill. We use a set of individual and household characteristics, distance to and quality of the nearest private and public health facility, and a dummy, which is equal to one when the nearest public facility is a pharmacy. This control is introduced because we are interested in allocations in the health budget, which does not consider pharmacies as primary health care facilities, even though households appear to be using them as such.[14] The first two models we estimate are simple probit models explaining usage among people who are ill, as well as the probit random effects model. The probit random effects model controls for constant individual-specific effects that are randomly distributed across the population, effectively allowing for serial dependence of errors across the same individuals in the panel.

We experimented by entering the three measures of quality in the usage regression both individually and as a composite. The composite is convenient because we can use it in later calculations of the cost of providing clinics of satisfactory quality relative to the benefits of increased usage. In building the composite we score the three attributes using the information in table 3. The equal weighting is arbitrary. However, when each of the three components was entered separately, the results were not substantially affected.[15]

The empirical model used is a binary choice model focusing on whether treatment is sought or not. This choice needs justification. It is more usual to run a multinominal or nested multinominal logit model to explain the choices between different types of facilities (as in Gertler and van der Gaag 1990 or Lavy and Germain 1994), thereby revealing the substitution effects between different facilities. However, our focus is on how total usage of modern health care is affected. We seek to quantify the net effect of changes in the quality of and distance to the primary public facilities of health care on total usage (including clinics, pharmacies, and hospitals but excluding traditional healers). Thus substitutions between facilities are automatically netted out from our results. We discuss the implications of using a multinominal choice model.

The results of these two models, using the composite quality index, are quite similar (table 4). The strongest effects are as follows: Primary education of the main female adult increases health usage and so does higher real consumption per capita. Distance to a public facility reduces usage, but the effect of distance to a private facility is insignificant. The public facilities' quality index is signifi-

14. An alternative would have been to include only data on public and private health facilities that are not pharmacies, but this distinction was not made during data collection. Qualitatively, results on the key variables of interest are unchanged when excluding this dummy, although the fit is not as good.

15 Testing the joint restriction of equality of the coefficients on the three quality indicators for the public facilities revealed that equality could not be rejected at 1 percent. Individually, we typically found that the presence of a nurse in a public facility had no effect on usage, whereas inputs and equipment (drugs and refrigerators) had strong significant effects.

TABLE 4. Regression Results Explaining Health Care Usage, 1994–97
(dependent variable = consulted modern facility when ill)

	Model (1) Probit (pooled cross-section)		Model (2) Random effects probit model (unbalanced panel 1994–97)	
Item	Coefficient	Probability > |z|	Coefficient	Probability > |z|
Constant	−1.184	0.000	−1.642	0.000
Age in months/1,000	0.546	0.263	0.763	0.096
Age squared/100,000	−0.767	0.107	−1.040	0.045
Sex	0.063	0.364	0.108	0.158
Sex of head of household	0.108	0.195	0.172	0.099
Age of head of household	−0.002	0.519	−0.003	0.327
Mother completed primary school	0.703	0.068	1.016	0.010
Household size	0.012	0.506	0.016	0.193
Log consumption per capita (real)	0.122	0.000	0.140	0.005
Kilometers to nearest public health care facility	−0.032	0.008	−0.039	0.000
Quality index public	0.383	0.069	0.539	0.000
Kilometers to nearest private health care facility	−0.001	0.944	−0.001	0.822
Quality index private	0.451	0.059	0.699	0.058
Dummy public pharmacy	0.581	0.003	0.873	0.011
Sample selection				
Joint significance	78.73	0.000	81.06	0.000
Sample size	2,317		2,317(1,834 groups)	

Note: Errors corrected for cluster effects.
Source: Ethiopian Rural Household Survey, rounds 1–4.

cantly different from zero, while not significantly different from the private quality effects. We experimented with adding more variables into the equation, without changing the findings. For example, it may be argued that quality and facility presence are likely to interact in the demand for health care; that is, distance and poor quality may reinforce each other. However, multiplicative terms for private and public facilities were found to be jointly and separately insignificant when entered into the models.[16]

Marginal effects of the key variables for these models (as for several others) are reported in table 5 (evaluated at the mean). The first two rows give the marginal effects for the models reported in table 4. Broadly speaking, they are quite similar. For example, despite significant sample selection terms, the specification with and without sample selection has very similar marginal effects.

Using model (2) we find that having a mother who completed at least primary school increases health care usage about 35 percentage points. A 10 per-

16. For example, adding public and private quality times distance terms in model 2 gave z-values of 0.57 and −0.82 and $\chi^2(2) = 1.43$ for their joint significance in the model (that is, no significant effects).

TABLE 5. Marginal Effects of Different Specifications, Evaluated at Mean
(dependent variable = use of modern health facility)

Model	Mother completed primary school	Log consumption per capita	km to nearest public facility	Quality index public facility	km to nearest private facility	Quality index private facility
Pooled cross-section probit model (see table 4, model [1]) (N = 2,317)	0.260 (0.065)	0.049 (0.001)	–0.013 (0.011)	0.153 (0.071)	0.000 (0.951)	0.180 (0.109)
Probit random effects model (see table 4, model 2) (N = 2,317)	0.347 (0.000)	0.056 (0.005)	–0.016 (0 000)	0.215 (0.000)	0.000 (0.822)	0.279 (0.058)
Probit sample selection model (N = 2,242)	0.244 (0.162)	0.035 (0.042)	–0.012 (0.170)	0.157 (0.352)	0.000 (0.714)	0.150 (0.031)
Probit random effects, considering pharmacy as no treatment (N = 2,317)	0.301 (0.011)	0.052 (0.002)	-0.017 (0.000)	0.172 (0.000)	0.000 (0.145)	0.182 (0.140)
Probit random effects (model [1]) plus knowledge variables (N = 2,242)	0.355 (0.027)	0.049 (0.015)	–0.015 (0.000)	0.190 (0.000)	0.000 (0.904)	0.229 (0.121)
Multinominal logit model, marginal probability of going for treatment (N = 2,194)	0.242 (0.000)	0.059 (0.000)	–0.016 (0.000)	0.200 (0.000)	0.000 (0.000)	0.267 (0.000)
Probit random effects model, excluding children below age 10 (N = 1,761)	0.246 (0.399)	0 050 (0.022)	–0.015 (0.001)	0.230 (0.000)	0 000 (0.878)	0.029 (0.868)
Probit random effects model, those below median per capita consumption only (N = 1,225)	0.344 (0.233)	0.116 (0.007)	–0.015 (0.008)	0.234 (0.002)	0.000 (0.827)	–0.032 (0.908)
Logit model, restricting sample to four panel villages (N = 719)	0.069 (0.777)	0.122 (0.000)	–0.030 (0.000)	0.352 (0.000)	0.003 (0.000)	–0.222 (0.000)
Fixed effect logit model (four villages) (N = 110)		–0.048 (0.378)		0.583 (0.226)		

Note All marginal effects were calculated as the percentage point usage effect of a change in the variable of interest by one on usage (dV/dx). For dummy variables (primary education effect), the effect is for a discrete change from 0 to 1.

Source· Authors' calculations

cent increase in consumption increases usage by almost 6 percentage points. A one-km closer public health facility increases use by 1.6 percentage points, whereas an increase in the quality of public facilities by a third (that is, one additional item in the index satisfied) increases usage about 5 percentage points. A similar effect applies to the quality of private facilities. The random effects formulation typically has somewhat higher estimates. However, in the rest of the article we are interested in the relative contribution of increasing usage by reducing the distance to or increasing the quality of health care facilities, and the relative marginal effects remain very similar across specifications.

The rest of table 5 looks into these marginal effects based on tests of the robustness of these results. The third row reports an attempt to control for possible self-selection problems associated with censoring the sample to those who are ill, by estimating a bivariate probit with sample selection with appropriate Heckman correction. We jointly estimate a selection equation and the usage equation, allowing for correlation in the errors and using this information in the usage equation as a control for selection (Greene 1993). The key issue is to find identifying instruments for the equation explaining self-reported illness. Although a reasonable expectation is that those households that are poorer and have less education would have a higher incidence of illness, data based on the self-reporting of illness do not typically find such a pattern. This is also the case here, suggesting that some of the problems related to self-reporting may have to do with information about illness, a fact that can be exploited to find appropriate instruments. In particular, we run a first-stage regression explaining illness using the same variables as in the usage regression plus variables describing health knowledge. In the survey the most important female household member (usually the spouse of the head or the head) was asked to identify the cause of diarrhea and malaria. The instruments used were variables based on whether these answers were correct or not, both directly entered into the regression and interacted with consumption and education variables. Knowledge as an identifying instrument may not be fully convincing, but even so, the robustness of the findings is striking.[17]

Row 4 displays the results when visits to pharmacies are not treated as modern health care visits; row 5 considers a specification including health knowledge variables (as in the selection equation in row 3). Both models give very similar marginal effects as before.[18] A multinominal specification (using a multinominal

17 As always in selection models, it could be argued that illness knowledge will also contribute to whether an ill person seeks treatment, making the usage model misspecified (if left out) or the selection equation identified only by distributional characteristics (if included in the usage equation). This problem cannot be solved. Furthermore, knowledge may be a function of past usage. If it is, this variable measures past usage. The high significance is therefore not surprising, and the variable is not very convincing as an identifying instrument.

18 The specification displayed in row 5 suggests that controlling for wealth, education, and facility characteristics, the return to health knowledge is high: Knowing the cause of diarrhea increases modern health usage 5 percent, while knowing the cause of malaria increases modern health usage 4 percent.

logit with choices of no treatment, treatment in private facilities, and treatment at public facilities) did not give different results. Restricting the sample to adults and investigating whether the poor behave differently also revealed little difference, especially in the variables related to public health facilities.

The last two rows warrant more discussion. A major potential shortcoming of this analysis is that quality and distance effects are measured without taking account of the fact that facilities may not have been randomly placed. In particular, planners may have located them according to characteristics unobserved by the researcher that are correlated with the determinants of health and health usage, biasing estimates on the variables of interest. To the extent that these unobserved characteristics are fixed, fixed effects models can deal with this problem. We used a fixed effects logit model on the four villages for which we have information. For comparison, we first gave the logit for this restricted dataset. The sample is very small to obtain good point estimates of the various marginal effects. In general, the logit model gives larger marginal effects on quality and distance of public facilities, although in relative terms the effects are still very similar. The fixed effects regression can give some indication of the bias involved as a result of program placement (or indeed any other missing fixed characteristic). The sample becomes restricted to people who were ill in at least 2 of the 4 periods and who experienced a change in health usage in this period. Only 110 observations satisfy these conditions, and only household size, consumption, and quality of public health facilities appear to change in this period. Quality changed in only 2 of the 4 villages, so this is indeed a small sample to pick up any effects.

In view of this, the results are quite interesting. We find a positive (albeit not quite significant) correlation between quality and usage. The effect was larger than the effect in the simple logit on the four villages (and outside the implied 95 percent confidence interval by this model). This suggests that the quality effect is underestimated due to placement effects; that is, an unobserved effect that contributes positively to usage is negatively correlated with the quality indicator. For example, if the unobserved effects are linked to wealth, then the supply rule of facility quality may have favored poorer areas. Because we do not observe new facility building (and therefore change in distance) in the sample, we cannot investigate further the placement effects related to distance. Even so, the results of the fixed effects model suggest that we may well be underestimating the contribution of quality to health usage, biasing the results in favor of allocating public expenditure toward reducing distance rather than increasing quality.

Thus both the distance and the quality aspects of the shadow price of usage have significant effects on demand. It is possible to compare these effects more directly. Restricting the discussion to the point estimates using the full sample, the marginal effect of increasing the index of quality from zero to one is 15–21 percent. A reduction in distance by one km increases usage by 1.2–1.7 percent. Obviously, these are only point estimates—for example, taking the marginal effects using the random effects model suggests a 95 percent confidence interval of about 12–31 percent for quality increases from zero to one, and a 0.8–2.3

percent increase of usage by reducing distance to public facilities by one km.[19] Another way of looking at the results is to consider the impact on usage of bringing all clinics up to the quality standard of "fully satisfactory."[20] The effect on usage would be considerable, with demand increasing 9.3 percentage points (using model [2]). The same effect on demand could be achieved by reducing the mean distance to clinics 5.9 km.[21] In the next section we compare the costs of increasing usage by improving quality and by reducing distance.

IV. Budgetary Costs and the Density-Quality Choice

The decision to build more facilities or improve the quality of existing facilities has been decentralized to local government in Ethiopia. Currently, about 43 percent of the health budget is being devoted to building more facilities, so this is probably the single most important choice in health care budgeting.[22] We now show how the results on usage can be combined with local information to yield a workable decision rule at the local level and show that it is likely to change budget decisions substantially.

Recall from condition (17) that one condition for an allocation to be optimal is that the marginal dollar spent on improving quality must have a smaller impact on usage than the marginal dollar spent on increasing the number of facilities. We consider these two marginal effects in turn.

The marginal effect of expenditure on quality improvement can be decomposed into two terms:

$$\frac{dV}{dE_q} = \frac{dQ}{dE_q} \cdot \frac{dV}{dQ} \tag{18}$$

where E_q is public expenditure on quality of provision.

In the previous section we measured Q on the six-point scale shown in table 3 and estimated the second term on the right-hand side, dV/dQ. Using informa-

19. It is possible to compare these results with Ghana. Lavy and Germain (1994) show that the net increase in use of (public and private) health facilities from bringing drugs, infrastructure, and service to "optimal" levels in public facilities was equal to 3 percentage points. Because quality appears to be substantially lower in the Ethiopian public health service, the effects appear comparable. Similarly, the effect of halving the distance to the nearest public health facility increased the usage of health facilities by 5 percentage points at the mean. In Ethiopia the effect of halving distance would be equivalent to about 4.1 percentage point.

20. We also looked at the effects for different quality items by introducing them in the regressions rather than the composite index. We found that the move to full drug availability would increase usage 3.6 percent and installing and repairing refrigerators in all public facilities would increase usage about 6.6 percent.

21. Other regressions give very similar relative effects Using the extreme point estimates of marginal effects implied by table 5 for the full sample, we found that increasing demand by bringing all public facilities up to full basic quality could also be obtained by reducing distance 3.8–7.7 km. The values implied by the confidence interval would suggest 2.2–16.9 km. These values can be used to investigate the robustness of the results in the next section.

22. Figures are for 1997–99, before actual capital expenditures (World Bank 2000).

tion at the regional level, it will usually be straightforward for a budget agency to estimate the first term, the cost of achieving a marginal quality improvement in the health facilities in a locality. Here we illustrate using a crude national-level approximation. We stress that our purpose is primarily illustrative: information at the level where such decisions are routinely made will always be superior to these national-level data.

Nationally, health costs are not available at the level of disaggregation used in table 3, forcing us to rely on inferences from the total cost of clinics. Even at this level of aggregation the data are problematic because not all components of expenditure are channelled through the budget. The total cost of providing the current level of primary health care quality in the public sector is the total recurrent expenditure on primary health care plus off-budget recurrent expenditure on drugs less revenue from user charges, which is largely in the form of charges for drugs. Recurrent expenditure on primary health care (that is, for health centers and clinics plus some allocation for health programs and management) in the 1994/95 budget was 138 million birr (World Bank 1995). However, in 1994 most of Ethiopia's drug supply was provided off-budget by donors. Although (remarkably) there is insufficient financial information to provide an accurate costing of this provision, we estimate that the cost of drugs net of recovered costs is approximately equal to the value of gross expenditure on primary care by the government. Hence we estimate the total gross cost of primary care at about 276 million birr (see Government of Ethiopia 1997).[23] This expenditure, distributed over the 2,010 primary health stations, purchases the current frequency distribution of quality of primary health care (so that the unit cost of a facility is 137,300 birr).

Using the six-point quality scale of table 3, which ranges from 0 to 1, the population-weighted mean level of quality is 0.60. For purposes of illustration we assume that the cost of health care quality is linear in this index. Thus raising average quality from 0.6 to 0.7 would require an increase in the gross cost of primary care of one-sixth, or 46 million birr. From the coefficient on quality in the random effects usage regression (model [2]), this increase in quality would raise usage by 2.6 percent.[24] Analogously, an increase in expenditure on quality improvement of 10 million birr per year would increase usage about 0.6 percent.

23. The data indicate that this is likely to be an overestimate. One report estimates the total value of drugs supplied in Ethiopia at about 360 million birr (Mengesha 1996). Using the right-hand-side data we find per capita expenditure on drugs of about 8 birr a year, equivalent to about 400 million birr a year—somewhat higher than Mengesha's estimate. Very few people are exempt from charges for drugs and treatment in rural Ethiopia. Data from the right-hand side indicate exemption rates at health stations and centers of less than 10 percent. Unpublished data from the sectoral investment office at the Ministry of Health suggest somewhat higher exemptions rates, but even those data show that just 12–27 percent of patients were exempt. At the current level of quality at health facilities, households are often dependent on private pharmacies for drugs. Taken together, this would support the view that off-budget financing of drugs is lower in net terms than the estimate used in the evaluation—that is, we overestimate the cost of quality.

24. Later we discuss the sensitivity of the results to using different estimates.

Making such a calculation at the national level is difficult; in any particular local situation it will be much easier. There will be a limited number of clinics, most of which will be of satisfactory quality, and a few (typically about a third) will have identifiable deficiencies, such as lack of trained staff or deficient equipment, for which the cost of rectification can reasonably be estimated. Hence the cost of quality improvement is not an intrinsically difficult number to estimate; it is simply problematic to arrive at a national average cost from national data.

We next consider the effect of an increase of 10,000 birr in expenditure on the provision of facilities. The marginal effectiveness of such expenditure can be decomposed into three terms:

$$\frac{dV}{dE_f} = \frac{dN}{dE_f} \cdot \frac{dD}{dN} \cdot \frac{dV}{dD} \tag{19}$$

where E_f is public expenditure on facility provision and N the number of facilities.

The first of the three terms in equation (19), the cost of an additional health unit, consists of the annual cost of running the unit and the amortized cost of its construction. We have already calculated the cost of running a unit at 137,300 birr. The amortized cost of construction can be inferred from government estimates of the construction cost of a primary health care complex (a set of five rural health stations or clinics, managed by a health center, usually located in an urban center). The unit cost of building such a complex, as budgeted in the government's health plan, is 1.5 million birr, or 300,000 birr per health station (Ministry of Health 1995). We convert this into an annualized cost assuming a discount rate of 10 percent, yielding an annual cost of 30,000 birr.[25] Hence the approximate total annual unit cost of an additional primary health facility is 167,300 birr.

We now turn to the second term, the relation between the number of facilities and the mean distance to a facility. In any particular local situation this is a simple matter. Given the location of existing facilities and the distribution of the population, there will be some site for a proposed facility that will maximize the reduction in the distance of potential users. The reduction in distance achieved by the proposed location is then a straightforward calculation. To illustrate using national data requires some simplifying approximations.

Suppose that the allocation decision in question is to build several new facilities to provide coverage to a district that has previously not been served and over which the population is evenly distributed. The decision problem is to determine the appropriate number of facilities, N. The district covers an area of C square km. Each facility will serve a catchment area C_f, and between them the facilities will serve the entire area, so that

25. The Ministry of Economic Development and Planning estimates the cost of building five health stations and a health center at almost 3.4 million birr. It uses much lower depreciation rates for buildings, so that the annualized cost is about 45,000 birr per health station. Building health stations without a supporting health center would be cheaper, at an annualized cost of about 20,000 birr.

$$NC_f = C. \tag{20}$$

Because the total area to be served is given, any change in the number of facilities must have an offsetting proportionate change in the catchment area:

$$dlnN + dlnC_f = 0. \tag{21}$$

We approximate C_f by the circular area πr^2, where r is the radius of the catchment area. In this case the mean distance of users to the facility, D, is determined as $2r/3$. The term dD/dN is more conveniently analyzed as an elasticity, $d\ln D/d\ln N$. This can be decomposed into two simple terms:

$$\frac{d\ln D}{d\ln N} = \frac{d\ln C_f}{d\ln N}\frac{d\ln D}{d\ln C_f} = \frac{d\ln \pi r^2}{d\ln N}\frac{d\ln(2r/3)}{d\ln \pi r^2} \tag{22}$$

From equation (21), the first term on the right-hand side of equation (22) is minus unity, and the second term is 0.5. Hence a 1 percent increase in the number of facilities reduces the mean distance to a facility by 0.5 percent.[26]

The third term in equation (19) was estimated in the usage regression. According to the random effects probit, a reduction in the mean distance to a facility of 0.1 km increases usage by 0.12 percent. Given the current mean distance to a facility, this is a reduction of about 1.7 percent. Such a reduction in the mean distance would require an increase in the density of facilities of 3.5 percent (that is, $1.017^2 - 1$).

The cost of a 3.5 percent increase in the number of facilities is 12 million birr a year. Hence an increase in expenditure on the provision of facilities of 10 million birr would increase usage about 0.1 percent.

Bringing the analysis together, an increase in expenditure of 10 million birr would raise usage about 0.6 percent if spent on quality improvement but only about 0.1 percent if spent on more facilities. Thus given the current allocation of the budget, quality improvement appears to be more effective in increasing usage than building additional facilities.[27] Yet we know from equation (17) that

26 To see this more intuitively, imagine the proposed facilities arranged on a grid pattern. To halve the average distance to a clinic, it is necessary to reduce the scale of the existing grid by half. This implies reducing the catchment area of each facility by a factor of four Hence a halving of distance requires a quadrupling of the number of facilities. A similar result holds for the hexagonal catchment areas, which Losch (1954) shows to be optimal Were the population and clinics strung out along a single line of road like beads on a string, the mean distance could be halved simply by doubling the number of clinics. However, about half of the rural population is living in villages that are not even connected by all-weather roads

27. The point that spending on quality appears more effective than spending on reducing distance remains valid if any of the alternative estimates discussed in section III is used. For example, using any combination of the other point estimates for the marginal effects, the impact from spending on quality is always at least twice as large as the impact from distance reduction. The 95 percent intervals of the point estimates of the models used do not change the conclusion, although the difference becomes smaller. For example, the intervals of the marginal effects based on the random effects model (model [2]) suggest that even the lowest marginal effect from quality and the highest marginal impact from density would imply that increasing usage by increasing density is still 50 percent more expensive than increasing usage by improving quality.

this cannot be optimal. Because improving health outcomes rather than usage is the policy objective, once expenditure is efficiently allocated a marginal increase in spending on quality improvement will increase usage by less than if the money were spent on more facilities. Hence nationally the current budget allocation appears to be too heavily skewed toward provision of facilities.

We have stressed that the national level data are primarily illustrative. The conclusion that the budget is currently misallocated indeed rests on heroic assumptions. Nevertheless, the analysis suggests that sizable gains might be achieved in the efficiency of health expenditure if local budget decisions applied the decision rule set out herein. The results suggest the limits of increasing health care usage through changes in the health budget. Both expenditure reallocation and expenditure increases appear to have relatively modest effects on usage. Maternal education and household income both affect usage quite substantially. Hence education and rural development policies may be even more important for health usage objectives than health care policy itself.

V. Conclusion

This article provides a decision rule for choosing between the quantity and the quality of health care provision that could be applied at the local level. Using national data, we show that current priorities are unlikely to be cost-effective. Based on analysis of a large-sample rural household survey, we find that household usage of health facilities is sensitive to the quality of health care as well as to the distance to the nearest public facility. Given the current budget allocation, it appears to be considerably less costly to increase usage by improving quality than by bringing facilities closer to people. Health outcomes rather than usage are the ultimate concern of policy, and quality improvement directly improves both health outcomes and usage. In an efficient budget allocation, expenditure on building facilities would be reduced to the point at which its effect on usage was greater than that of expenditure on quality improvement.

References

Akin, John S., C. Griffin, D. Guilkey, and B. Popkin. 1986. "The Demand for Primary Health Care Services in Bicol Region of the Philippines." *Economic Development and Cultural Change* 34(4):755–82.

Akin, John S., D. Guilkey, and H. Denton. 1995. "Quality of Services and Demand for Health Care in Nigeria: A Multinominal Probit Estimation." *Social Sciences and Medicine* 40(11):1527–37.

Alderman, Harold, and V. Lavy. 1996. "Household Responses to Public Health Services: Cost and Quality Tradeoffs." *World Bank Research Observer* 11(1):3–21.

Appleton, Simon. 1998. "The Impact of Public Services on Health Care and Illness: A Treatment Effects Model with Sample Selectivity." *Journal of African Economies* 7(1):1–33.

Behrman, Jere R., and A. Deolalikar. 1990. "Health and Nutrition." In H. Chenery and T. N. Srinivasan, *Handbook of Development Economics*. Vol. 1. Amsterdam: North Holland.

Central Statistical Authority. 1999. *WMS 3–98: Report*. Addis Ababa.

Dercon, Stefan. 2000. "Poverty and Deprivation in Ethiopia." Background paper for *Poverty Assessment of Ethiopia*. World Bank, Africa Region, Washington, D.C.

Dercon, Stefan, and P. Krishnan. 1998. "Changes in Poverty in Rural Ethiopia 1989–95." CSAE (Centre for the Study of African Economies) Working Paper WPS 98/7. Oxford University, CSAE.

———. 2000. "In Sickness and in Health: Risk Sharing within Households." *Journal of Political Economy* 103(4):688–727.

Dor, Avi, and J. Van der Gaag. 1988. "The Demand for Medical Care in Developing Countries: Quantity Rationing in Rural Côte d'Ivoire." LSMS Working Paper 35. World Bank, Development Research Group, Washington, D.C.

Filmer, Deon, J. Hammer, and L. Pritchett. 1997. "Health Policy in Poor Countries: Weak Links in the Chain." Policy Research Working Paper 1874. World Bank, Development Research Group, Poverty and Human Resources, Washington, D.C.

Gertler Paul, and J. Van der Gaag. 1990. *The Willingness to Pay for Medical Care: Evidence from Two Developing Countries*. Baltimore: Johns Hopkins University Press for the World Bank.

Government of Ethiopia. 1997. *Ethiopia Social Sector Review: Public Expenditure Review II*. Addis Ababa: Ministry of Finance.

Greene, W. H. 1993. *Econometric Analysis,* 2nd ed. New York: Macmillan.

Hotchkiss, D. 1993. "The Role of Quality in the Demand for Health Care in Cebu, Philippines." University of North Carolina, Department of Economics, Chapel Hill, N.C.

Lavy, Victor, and J.–M.Germain. 1994. "Quality and Cost in Health Care Choice in Developing Countries." LSMS Working Paper 105. World Bank, Development Research Group, Washington, D.C.

Lavy, Victor, J. Strauss, D. Thomas, and P. De Vreyer. 1996. "Quality of Health Care, Survival, and Health Outcomes in Ghana." *Journal of Health Economics* 15(3):333–57.

Litvack, Jennie I., and C. Bodart. 1993. "User Fees and Improved Quality of Health Care Equals Improved Access: Results of a Field Experiment in Cameroon." *Social Science and Medicine* 37(3):369–83.

Losch, August. 1954. *The Economics of Location*. New Haven, Conn.: Yale University Press.

Mengesha, Yishak. 1996. "Issues and Problems in the Production, Import, and Distribution of Pharmaceutical Products in Ethiopia." Background Paper for the Social Sector Review PER II. Ministry of Finance, Addis Ababa.

Ministry of Health, Ethiopia. 1995. "Health Sector Development Programme: A 20-Year Long-Term Plan Framework." Planning and Programming Department, Addis Ababa.

———. 1999. "Health and Health-Related Indicators 1991 E.C." Planning and Programming Department, Addis Ababa.

Mwabu, Germano, M. Ainsworth, and A. Nyamete. 1993. "Quality of Medical Care and Choice of Medical Treatment in Kenya: An Empirical Analysis." *Journal of Human Resources* 28(4):838–62.

Rosenzweig, Mark R., and K. Wolpin. 1986. "Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions." *American Economic Review* 76(3):470–82.

Strauss, John, and D. Thomas. 1995. "Modelling Human Resources." In J. Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics*. Vol. 3A. New York: Elsevier.

Thomas, Duncan, V. Lavy, and J. Strauss. 1996. "Public Policy and Anthropometric Outcomes in the Côte d'Ivoire." *Journal of Public Economics* 61(2):155–92.

World Bank. 1995. *Ethiopia Public Expenditure Review: Issues in Public Expenditure.* Vol. I. World Bank, Africa Region, Washington, D.C.

———. 1996. *African Development Indicators*. World Bank, Africa Region, Washington, D.C.

———. 2000. *Ethiopia Public Expenditure Review*, Vols. I and II. World Bank, Africa Region, Washington, D.C.

WHO (World Health Organization). 2000. *World Health Report*. Geneva.

THE WORLD BANK ECONOMIC REVIEW, VOL. 16, NO. 3 449–472

A Firm's-Eye View of Commercial Policy and Fiscal Reforms in Cameroon

Bernard Gauthier, Isidro Soloaga, and James Tybout

After decades of high trade restrictions, fiscal distortions, and currency overvaluation, Cameroon implemented important commercial and fiscal policy reforms in 1994. Almost simultaneously, a major devaluation cut the international price of Cameroon's currency in half. This article examines the effects of those reforms on the incentive structure faced by manufacturing firms. Did the reforms create a coherent new set of signals? Did they reduce dispersion in tax burdens? Was the net effect to stimulate the production of tradable goods? The results of applying a cost function decomposition to detailed firm-level panel data suggest that the reforms created clear new signals for manufacturers, as effective protection rates fell by 80 to 120 percentage points. In contrast, neither the tax reforms nor the devaluation had a major systematic effect on profit margins. The devaluation did shift relative prices dramatically in favor of exportable goods, causing exporters to grow relatively rapidly.

On gaining independence in 1960, Cameroon adopted an interventionist approach to industrialization and development. Its commercial policies kept import prices high, and its tax code selectively promoted certain firms and penalized others. These policies continued into the late 1980s and early 1990s, when the distortions they created were compounded by significant currency overvaluation in the Communauté Financière Africaine (CFA) zone, of which Cameroon is a member. In the face of crisis, the CFA countries agreed to devalue in 1994. Almost simultaneously, Cameroon implemented significant commercial policy reforms and attempted to level the playing field by reducing tax system inequalities.

This article examines the effects of these reforms on the incentive structure faced by manufacturing firms. Did they create a coherent new set of signals? Was their net effect to stimulate the production of tradable goods? Each of these issues

Bernard Gauthier is Professor of Economics at the École des Hautes Études Commerciales in Montreal. His e-mail address is bernard.gauthier@hec.ca. Isidro Soloaga is Professor of Economics at Universidad de las Américas-Puebla, Mexico. His e-mail address is isoloaga@mail.udlap.mx. James Tybout is Professor of Economics at Pennsylvania State University and a Research Associate of the National Bureau of Economic Research. His e-mail address is jtybout@psu.edu. The authors would like to thank Shanta Devarajan, Mark Gersovitz, and Thomas McCool for helpful discussions and Michel Sylvain for assistance. They are also grateful to three anonymous reviewers and François Bourguignon for many useful comments and suggestions. This article was funded by the World Bank and the Central African Customs and Economic Union (UDEAC).

DOI. 10 1093/wber/lhf005

is addressed using annual survey data collected by the Regional Program on Enterprise Development (RPED), along with product-specific prices and quantities subsequently collected from a subset of the RPED sample.[1]

The strength of the analysis lies in the data on which it is based. For each type of tax and tariff, the firms in the sample reported the amounts they paid before and after the reforms. Because we revisited the sample firms to collect price and quantity information on their major inputs and outputs, we are able to impute the effects of tariffs on input prices from official tariff schedules for firms that did not directly import the intermediate goods they used. We are also able to gauge the relative importance of each input and output to each producer. In sum, the data provide a far more detailed basis for inference than is typically available.[2]

To organize the analysis, we use a cost function decomposition. Fiscal and commercial policy reforms are treated as affecting the effective prices of inputs and outputs faced by firms; their net effects are then calculated in terms of the changes they induced in costs per unit revenue, firm by firm. Assuming that international trade determines the border prices of all inputs and outputs, the calculations capture all the effects of Cameroon's fiscal and commercial policy reforms on the incentive structure and firms' gross profit margins.[3]

By using a cost function approach rather than input-output tables, we allow for the possibility that firms are able to substitute away from inputs that become relatively expensive and toward inputs that become relatively cheap. Similarly, intrafirm substitutions among final products are recognized. Our effective protection figures therefore give a better measure of the true burden on producers than the traditional calculations.

I. Tax and Commercial Policy Reforms

Until 1994 the Cameroonian government relied heavily on selective tax and tariff exemptions to promote industrial development. This strategy began in 1960, when the country enacted an Investment Code to attract foreign capital and encourage import-substituting industrialization. It was also shaped by the 1964 Treaty of Brazzaville, which dictated a number of taxes and duties to be implemented in all Central African Customs and Economic Union (*Union*

1. More details on the RPED surveys in Cameroon may be found in Gauthier (1995). Information on the follow-up surveys is provided in Tybout and others (1997).

2. Standard effective protection measures are based on input-output matrices at the two-digit or three-digit level, in combination with tariff schedules or international price comparisons (see, for example, Balassa 1965). Many examples of this type of calculation and further references can be found in the seven volumes of country studies produced for the World Bank's "Liberalizing Foreign Trade" project Michaely, Papageorgiou, and Choksi (1991) summarize the main findings.

3. If this assumption is too strong, our calculations isolate only the direct effects of policy reforms on after-tax, after-tariff prices of inputs and outputs. The general equilibrium effects of the reforms on pretax, pretariff prices are not ignored; they are lumped in with all other residual factors, such as the exchange rate, that affect relative prices.

Douanière et Économique de l'Afrique Centrale; UDEAC) countries.[4] Following these events, the Cameroonian government layered on additional special tax schemes and exemptions. The cumulative effect was to create one of the most complex and unfair systems of taxes and duties in Sub-Saharan Africa.

Under this pre-reform regime, firms that did not enjoy access to any of the special UDEAC-wide or Cameroonian programs were subject to a variety of direct and sales taxes. Firms that imported intermediate goods were subject to four tariffs unless they had special status.[5] (Unlike most countries, Cameroon incorporated the equivalent of its domestic sales taxes into these tariffs.) The overall tariff structure was highly diversified, with rates ranging from 0 to 500 percent (World Bank 1995, appendix 6). The regime not only encouraged evasion but also provided considerable incentives for firms to seek special treatment from the tax authorities. Such treatment was available to manufacturers through a variety of mechanisms on a case-by-case basis. The appendix provides details on the direct and indirect taxes, tariffs, and special programs in effect in the prereform period.

Beginning in the mid-1980s, several adverse external shocks hastened the collapse of this policy regime. In 1985 the U.S. dollar, in which most primary commodity prices are denominated, depreciated sharply. Then oil prices (in dollars) fell in 1986, and the prices of other exported commodities—cocoa, coffee, and rubber—followed suit in the following year. The combined effect of these shocks was to induce massive fiscal deficits. Between 1985 and 1991, government's total revenues fell by 51 percent, largely because tariff and export tax revenues fell by 49 percent. The government financed this revenue shortfall by accumulating arrears with domestic suppliers and public servants, which led a banking sector crisis and a strong recessive effect on the economy. Despite the implementation of a structural adjustment program in 1989, per capita income had fallen in 1993 to *half* of its 1986 value.

Finally, in January 1994, the government began to dismantle this policy regime. In several decrees, it attempted to correct antitrade biases by increasing the importance of domestic taxes and reducing tariffs. It also attempted to reduce inequalities, distortions, corruption among administrators, and incentives for evasion.[6] These reforms were partly motivated by the urgent need to restore fiscal balance and to lay the foundation for long-term economic recovery. But in addition, they were designed to comply with conditionality in a World Bank Structural Adjustment Program, and to further the UDEAC objective of promoting regional economic integration.

4. UDEAC is composed of Cameroon, the Central African Republic, Chad, the Democratic Republic of Congo, Equatorial Guinea, and Gabon. It was formed in 1964 by the Treaty of Brazzaville.

5. Although a handful of agro-industrial imports remained covered by quantitative restrictions in 1993, almost all nontariff barriers had been phased out by that time as part of a World Bank structural adjustment program.

6. In 1994 more than 50 percent of the 200 firms interviewed in the RPED sample reported that they had not paid their full tax obligations the previous fiscal year (Gauthier and Gersovitz 1997).

The new policy regime included several components that affected external trade:

- The four types of tariffs were replaced by a unified single system known as the common external tariff (TEC), applicable to imports from non-UDEAC countries.
- Imports were classified into four categories, with tariff rates of 5–35 percent.
- A general preferential tariff was introduced for trade between UDEAC countries, with an initial rate fixed at 20 percent of the applicable TEC.[7]

The reform package also essentially replaced the various sales taxes with a value-added tax and eliminated most special privileges. (The appendix provides details.)

Table 1 documents the coverage of special fiscal regimes within the RPED sample before and after the reforms. The proportion of manufacturing enterprises enjoying fiscal privileges dropped from 65 percent to 14 percent over the two-year period. Measured in terms of sales or share of the total tax burden, the phase-out of privileges was equally dramatic. However, most of the major importing firms continued to enjoy special privileges after the reforms. Special regimes applied to 99 percent of the total value of sample imports in 1992/93 and to 74 percent in 1994/95. This pattern reflects the fact that major importers in Cameroon are large, and large firms continued to receive privileges.

If the reforms had bite, many firms that had enjoyed special status in 1992/93 should have borne a larger tax burden in 1994/95. To quantify this effect, we present the tax rates firms reported facing in each fiscal year (table 2). The firms are grouped into three categories: those in special programs (which lost most of their benefits), those with free trade zone status or ad hoc agreements (some of which retained their benefits), and firms operating under the common law regime in 1992/93.

Firms with special incentive programs in 1992/93 reported that they faced an average sales tax rate of 8.4 percent that year, whereas in 1994/95 they were confronted with an average quasi-value-added tax of 14.9 percent. Similar patterns emerge for the free trade zone/ad hoc agreement group and the unprivileged group, although their rates were generally not as favorable as those of the special program firms. The special program group enjoyed a discount of several percentage points, and there was no obvious tendency for this group to converge toward the others. Furthermore, as a percentage of sales, the 1994/95 value-added taxes were generally lower than the 1992/93 turnover tax. (See the 1994/95 figures in parentheses.) Thus, although it is possible that the tax burden was spread more evenly among the privileged firms after the reforms, it did not increase for them on average.

With respect to customs, the rates faced by the firms that originally enjoyed special programs increased from 15.8 percent in 1992/93 to 19.8 percent in 1994/

7. This rate was reduced to 10 percent on January 1, 1996, and eliminated altogether on January 1, 1998.

TABLE 1. Coverage of Special Regimes
(percent)

Item	1992/93	1993/94	1994/95
Percentage of firms enjoying at least one special tax regime	64.8 (83)	60.9 (78)	14.1 (18)
Privileged firms' sales as a percentage of total sales	98.5	94.4	29.5
Privileged firms' imports as a percentage of total imports	99.1	98.2	74.0
Privileged firms' taxes as a Percentage of total taxes	98.3	97.4	22.8

Note: Number of firms in each category is given in parentheses Total number of firms is 128
Source: Authors' calculations.

95, as privileges were phased out. The free trade zone firms and firms with ad hoc arrangements faced an even greater increase, with average rates jumping from 18.5 percent to 30 percent. This reflects the fact that more than half of the sample firms under these regimes lost their privileges after 1992/93. For firms operating under the normal regime in 1992/93, average customs rates fell from 66.8 percent in 1992/93 to 20.2 percent in 1994/95. Thus there is some evidence that the tariff reforms tended to level the playing field.

II. QUANTIFYING THE EFFECTS OF COMMERCIAL POLICY AND FISCAL REFORMS

The reforms did indeed change the level and distribution of the tax burden. However, the data do not document the combined effects of these reforms on after-tax costs per unit revenue for individual firms. This is our next objective. As in Tybout and others (1997), we begin with a cost function:

$$C = f\left(Q, \tilde{P}_L, \tilde{P}_I, \tilde{P}_K, A\right) \tag{1}$$

TABLE 2. Average Indirect Tax Rates for Different Categories of Firms Based on 1992/93 Status
(percent)

Item	1992/93	1993/94	1994/95
Sales or value-added taxes			
Special incentive programs (UDEAC and Cameroon)	8.4	8.3	14.9 (7.0)
Free trade zone or ad hoc agreements	10.9	8.7	16.5 (5.9)
No privileges	10.3	10.7	16.0 (9.5)
Customs			
Special incentive programs (UDEAC and Cameroon)	15.8	17.8	19.8
Free trade zone or ad hoc agreements	18.5	—	30.0
No privileges	66.8	52.4	20.2

— Not available.
Note. Total number of firms is 128 Figures are cross-firm averages of 1994/95 *taxes sur le chiffre d'affaires* (TCA; sales tax). Figures in parentheses are averages of 1994/95 TCA weighted by the ratio of value-added to total sales.
Source. Authors' calculations.

where C is the minimum attainable cost of producing output level Q, given productivity level A and the vector of effective (after tax, after tariff) prices for intermediate goods, $\tilde{P}_I$; labor, $\tilde{P}_L$; and capital, $\tilde{P}_K$. By Shephard's lemma, we have:

$$(2) \qquad d\ln C = \left(\frac{1}{\eta}\right) d\ln Q + s_I\left(d\ln \tilde{P}_I\right) + s_L\left(d\ln \tilde{P}_L\right) + s_K\left(d\ln \tilde{P}_K\right) + \left(\frac{\partial \ln C}{\partial \ln A}\right) d\ln A$$

where s_j denotes the share in total cost of the jth factor $\left(\sum_{j=1}^{J} s_j = 1\right)$ and η is the elasticity of output with respect to cost, or returns to scale. Normalizing by growth in the value of output, we obtain a decomposition of the sources of growth in cost per unit revenue:

$$(3) \qquad \begin{aligned} d\ln C - d\ln\left(Q\tilde{P}_Q\right) = \left(\frac{1}{\eta} - 1\right) d\ln Q + s_I\left(d\ln \tilde{P}_I - d\ln \tilde{P}_Q\right) + s_L\left(d\ln \tilde{P}_L - d\ln \tilde{P}_Q\right) \\ + s_K\left(d\ln \tilde{P}_K - d\ln \tilde{P}_Q\right) + \left(\frac{\partial \ln C}{\partial \ln A}\right) d\ln A \end{aligned}$$

(Note that unlike effective input prices, the effective output price, $\tilde{P}_Q$, is the pretax price to the buyer.) A second-order Tornqvist approximation to this expression in discrete time is given by:

$$(4) \qquad \begin{aligned} \Delta\ln C - \Delta\ln\left(Q\tilde{P}_Q\right) = \left(\frac{1}{\eta} - 1\right) \Delta\ln Q + \overline{s_I}\left(\Delta\ln \tilde{P}_I - \Delta\ln \tilde{P}_Q\right) + \overline{s_L}\left(\Delta\ln \tilde{P}_L - \Delta\ln \tilde{P}_Q\right) \\ + \overline{s_K}\left(\Delta\ln \tilde{P}_K - \Delta\ln \tilde{P}_Q\right) + \left(\frac{\partial \ln C}{\partial \ln A}\right) d\ln A \end{aligned}$$

where Δ is the difference operator for period t versus $t-1$ and overbars denote cross-period averages of the associated variable.

Commercial policy affects costs per unit revenue by changing the after-tariff prices of inputs and outputs. Domestic tax policy similarly affects input and output prices net of taxes and may further change after-tax costs through lump sum taxes such as the *patente* (see appendix). The rest of this article is devoted to quantifying these channels of transmission from policy reforms to the incentive structure at the firm level.

It is possible that commercial and domestic tax policy affect the efficiency parameter, A. Similarly, if there are scale economies, they may affect unit costs by changing the volume of output.[8] These channels of transmission are empirically intractable, and we do not attempt to measure their separate effects on A.

Perhaps more important are some dimensions of response that our methodology misses entirely. In particular, some relatively efficient firms were presum-

8 Head and Reis (1999) provide a recent survey of the theoretical channels through which commercial policy can affect scale efficiency.

ably able to enter because government connections became less important, and some relatively inefficient firms were presumably forced to exit for the same reason. Without data on the population of plants, however, we cannot measure the net benefits generated by induced entry or exit.

Linking Prices to Policy

Suppose for the moment that every good used or produced by Cameroon firms is also available in foreign markets and that arbitrage between domestic and foreign goods is perfect. It is then straightforward to calculate the effects of the fiscal and commercial policy on the after-tax, after-tariff prices faced by producers.

Specifically, under the prereform regime, directly imported inputs were subject to tariffs but not sales taxes (t), and domestically produced inputs were subject to sales taxes but not tariffs. With perfect arbitrage, Cameroon firms paid $\tilde{P}_{I_i} = P_{I_i}(1 + t) = P^*_{I_i}(1 + \tau_{I_i})$ for the ith input, where $P^*_{I_i}$ is the external price of this input, P_{I_i} is the pretax price of the domestically produced version of input i, τ_{I_i} is the tariff rate, and t is the sales tax rate. Analogously, after taxes a Cameroon producer of the jth output received $\tilde{P}_{Q_j} = P_{Q_j} = P^*_{Q_j}(1 + \tau_{Q_j}) / (1 + t)$ per unit produced.

When Cameroon moved to a value-added tax, domestic and foreign purchases of the ith input were effectively tax free (albeit not tariff free) because the value-added taxes paid on these purchases were rebated. But perfect arbitrage implies that the price of domestic inputs still matched the tariff-distorted world price, $\tilde{P}_{I_i} = P^*_{I_i}(1 + \tau_{I_i})$. Hence under the perfect arbitrage assumption, Cameroon's fiscal and commercial policy reforms influenced input prices only by affecting tariff rates. In contrast, in the product markets the new regime meant that Cameroon firms collected the tariff-distorted world price adjusted upward by the value-added tax rate (v), $P_{Q_j}(1 + t) = P^*_{Q_j}(1 + \tau_{Q_j})(1 + v)$, and kept $\tilde{P}_{Q_j} = P^*_{Q_j}(1 + \tau_{Q_j})$. Under the perfect arbitrage assumption, then, moving to a value-added system increased the after-tax price of outputs relative to inputs by eliminating the cascading effect of sales taxes.

Of course, perfect arbitrage is not a realistic assumption for most products. Transaction costs and product differentiation will typically allow domestic and foreign varieties of the same good to command different prices, and the response of these prices to changes in commercial policy and the fiscal regime will doubtless depend on firm-specific perceptions of demand elasticities, if not strategic considerations. Dealing properly with these problems would require an extremely detailed computable general equilibrium model. No such models exist for Cameroon, nor is it feasible to construct one.

Because the general equilibrium and mark-up effects are too complex to disentangle, we isolate the discrepancy between domestic and foreign prices in the endogenous scaling variables, λ_{I_i}, which applies to the ith input, and λ_{Q_j}, which applies to the jth output. Accordingly, the effective price of the ith domestic input

TABLE 3. Effective Producer Prices ($\tilde{P}_I$, $\tilde{P}_Q$) under Alternative Regimes

Item	VAT regime		Sales tax regime	
Outputs ($\tilde{P}_Q$)	$\lambda_Q P_Q^* (1 + \tau_Q)$		$\lambda_Q P_Q^* (1 + \tau_Q) / (1 + t)$	
	Domestic	Imported	Domestic	Imported
Inputs ($\tilde{P}_I$)	$\lambda_I P_I^* (1 + \tau_I)$	$P_I^* (1 + \tau_I)$	$\lambda_I P_I^* (1 + \tau_I)$	$P_I^* (1 + \tau_I)$

Note. Input prices with tildas are inclusive of taxes and tariffs; output prices with tilde are exclusive of any taxes collected and passed on to the government. Prices with asterisks are pretax border prices, converted to domestic currency.

is $\tilde{P}_{Ii} = P_{Ii}^* \lambda_{Ii} (1+ \tau_{Ii})$ and the price of the jth domestically produced output is either $\tilde{P}_{Qj} = \lambda_{Qj} P_{Qj}^* (1 + \tau_{Qj}) / (1 + t)$ or $\tilde{P}_{Qj} = \lambda_{Qj} P_{Qj}^* (1 + \tau_{Qj})$, depending on whether the old or the new regime is in force. These relationships are summarized in table 3.

Before we substitute these producer prices back into equation (4), we must deal with the fact that firms use multiple inputs and produce multiple outputs. We use Tornqvist indices of the growth rates in effective input and output prices, which amount to share-weighted aggregations of the growth rates in the prices of the individual goods. Specifically, for intermediate inputs, we calculate

$$(5)\qquad \begin{aligned} \Delta \ln \tilde{P}_I &= \sum_{i=1}^{N} \bar{s}_i \Delta \ln \tilde{P}_{Ii} = \sum_{i=1}^{N} \bar{s}_i \Delta \ln\left(\tilde{P}_{Ii}^*\right) + \sum_{i=1}^{N} \bar{s}_i \Delta \ln(1+\tau_{Ii}) + \sum_{i=1}^{N} \bar{s}_i \Delta \ln\left(\lambda_I'\right) \\ &= \Delta \ln \tilde{P}_I^* + \Delta \ln(1+\tau_I) + \Delta \ln(\lambda_I) \end{aligned}$$

where $\bar{s}_i$ is the share of expenditures on the ith input (inclusive of tariffs) in total intermediate input costs, averaged across periods. Given that producers report prices paid inclusive of tariffs, as well as tariffs paid, we observe both $\tilde{P}_{Ii}$'s and τ_{Ii}'s, so the left-hand side and the tariff component of the right-hand side can be isolated. However, we do not have micro-data on the external prices of each product, so we cannot disaggregate the sum $\Delta \ln P_I^* + \Delta \ln(\lambda_I)$.

Analogously, for effective output prices we write:

$$(6)\qquad \begin{aligned} \Delta \ln\left(\tilde{P}_Q\right) &= \Delta \ln(1+\tau_Q) + \Delta \ln\left(P_Q^*\right) - \Delta \ln(1+t_Q) + \Delta \ln(\lambda_Q) \\ &= \sum_{j=1}^{J} \bar{\alpha}_j \Delta \ln(1+\tau_Q) + \sum_{j=1}^{J} \bar{\alpha}_j \Delta \ln\left(P_{Qj}^*\right) - \sum_{j=1}^{J} \bar{\alpha}_j \Delta \ln(1+t_{Qj}) + \sum_{j=1}^{J} \bar{\alpha}_j \Delta \ln(\lambda_{Qj}) \end{aligned}$$

where $\bar{\alpha}_j$ is the average share of the jth product in total revenues in periods t and $t - 1$. It should be remembered from table 3 that the sales tax is phased out between the initial and the final period, so $\Delta \ln(1 + t_Q)$ amounts to $-\ln(1 + t_Q^0)$, where t_Q^0 is the prereform sales tax rate. Also, as with effective input prices, it is not possible to distinguish the effects of imperfect arbitrage from the effects of changes in external prices.

A Generalized Cost Decomposition

Substituting these relative price expressions into the unit cost decomposition (equation [4]) and writing costs and revenues as net of taxes yields:

$$
\begin{aligned}
(7) \qquad \Delta \ln C - \Delta \ln\left(Q\tilde{P}_Q\right) = & \left(\frac{1}{\eta} - 1\right)\Delta \ln Q + \left(\frac{\partial \ln C}{\partial \ln A}\right)\Delta \ln A \\
& + \overline{s_I}\Delta \ln\left(1+\tau_I\right) - \Delta \ln\left(1+\tau_Q\right) \\
& + \Delta \ln\left(1+t_Q\right) \\
& + \overline{s_I}\left[\Delta \ln\left(P_I^*\lambda_I\right) - \Delta \ln\left(P_Q^*\lambda_Q\right)\right] \\
& + \overline{s_L}\left[\Delta \ln \tilde{P}_L - \Delta \ln\left(P_Q^*\lambda_Q\right)\right] + \overline{s_K}\left[\Delta \ln \tilde{P}_K - \Delta \ln\left(P_Q^*\lambda_Q\right)\right]
\end{aligned}
$$

The first line on the right-hand side reflects the scale and other efficiency effects, which we treat as a residual; the second line reflects the direct effects of commercial policy on unit costs; the third line reflects the direct effect of eliminating sales taxes; and the last two lines reflect the changes in relative prices not directly related to commercial policy or taxes. Of course, the general equilibrium effects of these policy changes come partly through λ_I, λ_Q, $\tilde{P}_L$, and $\tilde{P}_K$. We are unable to isolate these indirect effects. Note also that under the perfect arbitrage assumption (that is, when $\lambda_I = \lambda_Q = 1$), these last lines simply pick up changes in wages and world prices.

Because we are unable to observe effective prices for capital services directly, we henceforth assume that they grow at the same rate as the pretariff rate of growth in domestic output prices, $\Delta \ln(P_Q^* \lambda_Q)$. The last line then becomes a wage effect alone:

$$
\begin{aligned}
(7') \; \Delta \ln C - \Delta \ln\left(Q\tilde{P}_Q\right) = & \left(\frac{1}{\eta} - 1\right)\Delta \ln Q + \left(\frac{\partial \ln C}{\partial \ln A}\right)\Delta \ln A && \textit{(residual efficiency effect)} \\
& + \overline{s_I}\Delta \ln\left(1+\tau_I\right) - \Delta \ln\left(1+\tau_Q\right) && \textit{(effective protection effect)} \\
& + \Delta \ln\left(1+t_Q\right) && \textit{(tax reform effect)} \\
& + \overline{s_I}\left[\Delta \ln\left(P_I^*\lambda_I\right) - \Delta \ln\left(P_Q^*\lambda_Q\right)\right] && \textit{(relative pre-tax input price effect)} \\
& + \overline{s_L}\left[\Delta \ln \tilde{P}_L - \Delta \ln\left(P_Q^*\lambda_Q\right)\right] && \textit{(relative cost of labor effect)}
\end{aligned}
$$

It is worth noting that equation (7) deals only with changes in marginal tax rates and misses the effects of lump sum taxes entirely. We experimented with a more general formula that accommodates lump sum taxes and found that they played a negligible role during the sample period.[9]

Measuring Dispersion in Protection

A major objective of the Cameroonian reforms was to reduce cross-firm dispersion in protection. To quantify the government's success in this regard, we need to measure the effects of protection on firm-specific unit cost levels rather than unit cost growth rates. For this purpose we use our decomposition to measure

9. To treat lump-sum taxes, define these taxes to be T and write costs inclusive of lump-sum taxes as $C^* = C + T$. The decomposition can then be generalized to $\Delta\ln(C^*) - \Delta\ln(QP_Q) = \theta[\Delta\ln(C) - \Delta\ln(QP_Q)] + (1-\theta)[\Delta\ln(T) - \Delta\ln(QP_Q)]$, where $\theta = C/(C+T)$ is the share of costs before lump-sum taxes in total costs and an overbar denotes the cross-period average. The first right-hand term is simply equation 7 weighted by θ, the second term picks up the effect of growth in the ratio of lump sum taxes to net revenue. We implemented this generalized decomposition on our data and found extremely small values for the second term.

the change in unit costs that would have occurred for each firm in moving from a hypothetical regime of zero tariffs to the tariff rates it actually paid. Cross-firm dispersion in this rate of unit cost increase—before versus after commercial policy reforms—provides a basis for assessing changes in the amount of preferential treatment in the tariff code.

Constructing these measures of net tariff protection requires several additional assumptions. First, in the tradition of most effective protection calculations, we assume perfect international arbitrage and set $\lambda_I = \lambda_Q = 1$. Second, we need figures for the hypothetical expenditure shares that would have prevailed if producers had faced zero tariffs. Our solution is to assume that the elasticity of substitution among all intermediate inputs is unity. Then the same shares prevail with and without tariffs, and the level form of the tariff effect in the second line of equation (7') becomes approximately $\tau_Q - s_I\tau_I = \tau_Q - \sum_{i=1}^{N} s_{Ii}\tau_{Ii}$.[10] This expression is a variant of the standard effective protection measure when expressed as a ratio to value-added per unit revenue:

$$\frac{\tau_Q - \sum_{i=1}^{N} s_i\tau_{Ii}}{1 - \sum_{i=1}^{N} s_i}.$$ [11]

III. The Data

The RPED surveys collected data on costs, sales, taxes, tariffs, and other variables from about 200 Cameroon firms for the fiscal years 1992/93 and 1994/95.[12] However, these surveys did not collect information on the prices of inputs and outputs. About 80 firms in the RPED data base were revisited as part of a recently completed project and asked for recall information on the values and quantities of their five major inputs and five major outputs in both fiscal years. Only 36 firms were able to supply complete and credible information, a subsample we henceforth refer to as the resurveyed firms.

Using this subsample, we constructed unit prices for each product by dividing the value of production by the number of units produced. For example, indexing products by j, we obtained, $P_{jt} = V_{jt} / Q_{jt}$, $j = 1, J$. Intermediate input prices and the cost of labor were imputed analogously. The prices were reported inclu-

10. This follows because $\ln(1 + x) \cong$ for small values.

11. The most common alternative approach is to assume there are no substitution possibilities at all among intermediate inputs. This approach implies that our translog cost function is a poor approximation to technology and that effective protection calculations are best done using input shares based on international prices.

12. The firms in these surveys do not constitute a stratified random sample because no sampling frame was available to the survey designers. Instead, firms were selected from the 1989 Directory of Businesses published by the Chamber of Commerce, as well as from business associations and cooperatives. They were chosen to be broadly representative of the size distribution across the four manufacturing sectors studied: textile and garments, wood products and furniture, food processing, and metal product and machinery. Gauthier (1995) provides further details.

sive of tariffs and sales taxes, so they correspond to the effective prices $\tilde{P}_{Q_t}$ and $\tilde{P}_{I_t}$ described earlier. We augmented tariff data reported by the firms with official tariff information by product line obtained from the Cameroon government. Hence we were able to impute $\Delta\ln(\lambda_Q P_Q^*)$ and $\Delta\ln(\lambda_I P_I^*)$ using the identities in Table 3.[13] Finally, with these building blocks, we were able to solve for the residual scale economy and productivity effect, $\left(\frac{1}{\eta}-1\right)\Delta\ln Q+\left(\frac{\partial\ln C}{\partial\ln A}\right)\Delta\ln A.$

Before reform tariffs (T) included the *droits de douanes* (DD), *droits d'entrée* (DE), *taxe sur le chiffre d'affaires à l'importation*, and *taxe complémentaire à l'importation* tariffs applied to firms operating under the normal regime and the *tax unique* (TU) or *taxe intérieure à la production* (TIP) applicable to imports for firms receiving special privileges. (The appendix provides descriptions of these tariffs and taxes.) After reforms tariffs included the TEC or *tarif préférentiel généralisé* (TPG). The tax burden (t) included the *impôt sur le chiffre d'affaires intérieur* (ICAI) for firms operating under the normal regime before the reform and the TU or TIP applicable to local sales for firms operating under a special regime. After reform the indirect tax burden is composed of the TCA. Further discussion of the data may be found in Tybout and others (1997).

IV. Basic Findings: Pooled Sample

Let us begin with an overview of the magnitudes of the different shocks to unit cost. Equation (7') provides the relevant decomposition; it is empirically rendered in table 4. We also report real output growth. Each mean component of our decomposition is accompanied by a t-ratio; asterisks indicate whether the means are significantly different from zero. (Tests are done under the assumption that the firm-specific realizations are independent and normally distributed.) Alternative renderings of the same decomposition based on output-weighted averages and medians are shown in tables 5 and 6. Medians are calculated component by component, so they do not satisfy our identity exactly. Table 7 provides descriptive statistics on the prices that are used to construct our unit cost decomposition. Table 8 reports the levels and dispersion in effective protection measures discussed in section III.

For the pooled sample of 36 firms, the average increase in unit costs was 8 percent and not significantly different from zero. But this mild cost increase reflected several more dramatic offsetting forces. The single most important shock was commercial policy reforms, which drove up cost per unit revenue by 20.5 percent (t-ratio 8.45) on average. Increases in the (pretariff) relative price of intermediate goods added an additional 5.5 percent (t-ratio 1.15). Offsetting these effects were tax reforms, which reduced unit costs 2.7 percent (t-ratio 5.4); re-

13 An interesting extension would be to exploit data on international prices and isolate growth in λ's from growth in P^*'s.

TABLE 4. Commercial Policy, Tax Reform, and Unit Production Costs (equation [7]): Unweighted Averages

Subsample (number of firms)	Net unit cost growth	Tariff effect, outputs (1)	Tariff effect, inputs (2)	Effective protection effect (1) + (2)	Labor price effects	Intermediate input price effects	Residual productivity effects	Domestic tax effects	Real output growth
Food (14)	0.081	0.176**	−0.027**	0.149**	−0.065	0.130**	−0.107	−0.026**	0.028
	(0.463)	(3.943)	(−2.349)	(3.241)	(−1.398)	(2.547)	(−0.746)	(−3.243)	(0.183)
Textiles (9)	0.029	0.243**	0.008	0.250**	−0.059	−0.029	−0.104	−0.029**	0.222
	(0.219)	(10.125)	(1.000)	(11.538)	(−0.932)	(−0.422)	(−0.429)	(−2.900)	(0.915)
Wood products (4)	0.230	0.355**	0.000	0.355**	0.012	−0.095	−0.034	−0.007	0.004
	(1.247)	(1775.0)	n.a.	(1775.0)	(0.289)	(−1.284)	(−0.301)	(−0.933)	(0.035)
Metal products (9)	0.063	0.216**	−0.037*	0.205**	−0.118*	0.091	−0.085	−0.027*	−0.335
	(0.604)	(5.143)	(−2.921)	(4.184)	(−1.914)	(1.162)	(−0.399)	(−2.250)	(−1.573)
Small (17)	0.053	0.239	−0.020**	0.219**	−0.053	0.077	−0.210*	−0.021**	0.084
	(0.353)	(1.060)	(−2.425)	(8.361)	(−1.316)	(1.470)	(−1.941)	(−2.793)	(0.430)
Medium (11)	0.045	0.209**	−0.016	0.193**	−0.117	0.004	0.000	−0.034**	−0.139
	(0.450)	(4.126)	(−1.561)	(3.903)	(−1.748)	(0.061)	(0.000)	(−3.317)	(−1.155)
Large (8)	0.185	0.208**	−0.016	0.193**	−0.032	0.081	−0.025	−0.031**	−0.063
	(1.553)	(3.440)	(−0.823)	(2.689)	(−1.052)	(1.076)	(−0.305)	(−3.812)	(−0.459)
All firms (36)	0.080	0.222**	−0.018	0.205**	−0.068**	0.055	−0.085	−0.027**	−0.017
	(0.996)	(9.867)	(−1.340)	(8.425)	(−2.386)	(1.150)	(−1.275)	(−5.400)	(−0.165)
Domestic input-intensive (18)	0.104	0.238**	−0.005	0.233**	−0.032	−0.030	−0.045	−0.022**	0.194
	(1.134)	(10.410)	(−0.707)	(10.629)	(−0 930)	(−0.957)	(−0.622)	(−3.457)	(1.556)
Imported input-intensive (18)	0.056	0.208**	−0.031**	0.177*	−0.104**	0.140**	−0.125	−0.033**	−0.227
	(0.416)	(5.316)	(−3.131)	(4.081)	(−2.322)	(2.434)	(−0.116)	(−4.243)	(−1.498)
Nonexporters 92/93 (24)	0.089	0.214**	−0.021**	0.192**	−0.056*	0.091**	−0.116	−0.021**	0.043
	(0.787)	(7.280)	(−2.939)	(6.030)	(−1.533)	(2.093)	(−1.257)	(−3.319)	(0.303)
Exporters 92/93 (12)	0.061	0.241**	−0.010	0.231**	−0.091*	−0.016	−0.023	−0 039**	−0.136
	(0.673)	(7.075)	(−0.770)	(6.351)	(−1.970)	(−0.280)	(−0.310)	(−5.004)	(−1.096)
Nonexporters 94/95 (21)	0.133	0.198**	−0.019**	0.179**	−0.060	0.048	−0.009	−0.025**	−0.177
	(1.064)	(6.771)	(−2.353)	(5.327)	(−1.440)	(0.944)	(−0.098)	(−3.819)	(−1.382)
Exporters 94/95 (15)	0.005	0.258**	−0.016	0.242**	−0.080**	0.065	−0.191**	−0.031**	0.208
	(0.060)	(7.513)	(−1.511)	(7.100)	(−2.152)	(1.361)	(−2.169)	(−4.002)	(1.323)

n.a. Not applicable.
*Significantly different from zero at the 90 percent confidence level.
**Significantly different from zero at the 95 percent confidence level.
Note. *t*-statistics are in parentheses.
Source: Authors' calculations.

TABLE 5. Commercial Policy, Tax Reform, and Unit Production Costs (Equation 7'): Output-Weighted Averages

Subsample (number of firms)	Net unit cost growth	Tariff effect, outputs (1)	Tariff effect, inputs (2)	Effective protection effect (1) + (2)	Labor price effects	Intermediate input price effects	Residual productivity effects	Domestic tax effects	Real output growth
Food (14)	0.300	0.209	−0.059	0.150	−0.034	0.160	0.064	−0.040	−0.075
Textiles (9)	−0.070	0.269	0.009	0.278	−0.059	−0.045	−0.224	−0.021	0.145
Wood products (4)	0.080	0.355	0.000	0.355	0.005	−0.112	−0.160	−0.008	0.091
Metal products (9)	0.018	0.378	−0.011	0.367	−0.140	−0.103	−0.075	−0.030	−0.242
Domestic input-intensive (18)	0.009	0.264	−0.017	0.247	−0.066	−0.021	−0.126	−0.025	0.226
Imported input-intensive (18)	0.161	0.286	−0.023	0.263	−0.074	−0.050	−0.043	−0 035	−0.297
Nonexporters 92/93 (24)	0.089	0.214	−0.021	0.193	−0.050	0.280	−0.116	−0.021	0.043
Exporters 92/93 (12)	0.061	0.241	−0.010	0.231	−0.075	−0.053	−0.023	−0.039	−0.137
Nonexporters 94/95 (21)	0.253	0.125	−0.035	0.091	−0.016	0.128	0.087	−0.033	−0.196
Exporters 94/95 (15)	0.082	0.274	−0.020	0.254	−0.090	−0.027	−0.086	−0.030	0.034
Small (17)	0.204	0.175	−0.038	0.137	−0.033	0.197	−0.067	−0.030	−0.261
Medium (11)	−0.034	0.285	−0.013	0.272	−0.116	−0.039	−0.106	−0.045	−0.052
Large (8)	0.103	0.277	−0.020	0.257	−0.061	0.016	−0.082	−0.026	−0.009
Total (36)	0.082	0.274	−0.020	0.254	−0.070	−0.013	−0.086	−0.030	−0.027

Source: Authors' calculations.

TABLE 6. Commercial Policy, Tax Reform, and Unit Production Costs (Equation [7']): Medians

Subsample (number of firms)	Net unit cost growth	Tariff effect, outputs (1)	Tariff effect, inputs (2)	Effective protection effect (1) + (2)	Labor price effects	Intermediate input price effects	Residual productivity effects	Domestic tax effects	Real output growth
Food (14)	0.343	0.209	–0.007	0.202	–0.011	0.123	0.067	–0.014	–0.120
Textiles (9)	0.143	0.244	0.000	0.244	–0.038	–0.079	–0.136	–0.027	0.117
Wood products (4)	0.130	0.356	0.000	0.356	–0.008	–0.078	–0.147	0.000	0.032
Metal products (9)	0.205	0.213	–0.030	0.180	–0.068	0.099	–0.096	–0.024	–0.315
Domestic input–intensive (18)	0.156	0.244	–0.000	0.238	–0.017	–0.013	–0.140	–0.009	0.083
Imported input–intensive (18)	0.259	0.224	–0.024	0.214	–0.055	0.151	–0.314	–0.021	–0.314
Nonexporters 92/93 (24)	0.263	0.224	–0.005	0.215	–0.020	0.077	0.005	–0.006	–0.028
Exporters 92/93 (12)	0.022	0.271	–0.001	0.262	–0.068	–0.023	–0.116	–0.042	–0.151
Nonexporters 94/95 (21)	0.324	0.215	–0.000	0.215	–0.013	0.045	0.083	–0.010	–0.298
Exporters 94/95 (15)	–0.018	0.266	–0.005	0.263	–0 089	0.048	–0.157	–0.024	0.049
Small (17)	0.205	0.233	0.000	0.215	–0 031	0.055	–0.139	–0.004	0.018
Medium (11)	0.143	0.244	–0 008	0.244	–0.046	0.017	–0.083	–0.027	–0.128
Large (8)	0.259	0.243	–0.004	0.246	–0.011	0.050	–0.035	– 0.027	–0.162
Total (36)	0.190	0.237	–0.002	0.230	–0.035	0.046	–0.067	–0.014	–0.116

Source: Authors' calculations.

TABLE 7. Growth in Prices of Output, Intermediate Input, and Labor among Resurveyed Subsample, 1992/93–1994/95
(cumulative percentages)

Item	Mean (x)	SD (s_x)	SD of mean ($s_x / \sqrt{n}$)	Median	Interquartile range
Pooled sample (36)					
Output price (P_Q)	37.3	67.3	11.2	21.4	2.0 to 51.9
Input price (P_I)	73 0	68.3	11.4	72.1	20.6 to 108.4
Wage rate (P_L)	21.4	65.2	10.9	11.7	-21.1 to 45.8
Relative input price (P_I / P_Q)	44.2	65.3	10.9	35.8	2.7 to 73.7
Relative labor cost (P_L / P_Q)	5.3	68.1	11.4	-19.3	-34.9 to 39.8
Domestic input-intensive (18)					
Output price (P_Q)	26.8	72.3	17.0	14.6	-8.6 to 37.4
Input price (P_I)	50.9	59.4	14 0	48.5	0.0 to 73.5
Wage rate (P_L)	37.3	75 4	17.8	33.3	-17.5 to 53.6
Relative input price (P_I / P_Q)	37.3	57.6	13.6	33.3	6.0 to56.0
Relative labor cost (P_L / P_Q)	31.9	80.2	18.9	21.3	-29.1 to 70.2
Imported input-intensive (18)					
Output price (P_Q)	47.9	62.2	14.7	33.2	19.2 to 58.5
Input price (P_I)	95.0	71.1	16.8	94.6	71.3 to 133.6
Wage rate (P_L)	5.4	50.3	11.9	3.1	-34.6 to 23.8
Relative input price (P_I / P_Q)	51.0	73.3	17.3	44.3	-2.2 to 98.3
Relative labor cost (P_L / P_Q)	-21.3	40.0	9.4	-31.8	-43.0 to -4.4
Nonexporters (24)					
Output price (P_Q)	25.3	56.7	11.6	18.5	-4.3 to 39.3
Input price (P_I)	73.2	71.5	14.6	69.6	10.7 to 120.5
Wage rate (P_L)	27.9	76.5	15.6	21.6	-32.3 to 57.0
Relative input price (P_I / P_Q)	50.7	61.9	12.6	44.3	6.5 to 91.1
Relative labor cost (P_L / P_Q)	15.8	72.0	14.7	4.7	-34.4 to 53.3
Exporters (12)					
Output price (P_Q)	61.4	82 2	23.7	50.6	15.0 to 84.2
Input price (P_I)	72.5	64.6	18.6	73.2	35.0 to 96 3
Wage rate (P_L)	8.3	32 2	9.3	3.3	-13.5 to 31.7
Relative input price (P_I / P_Q)	31.2	72.8	21.0	25.9	-8.7 to 52.9
Relative labor cost (P_L / P_Q)	-15 8	56.4	16.3	-30.7	-37 0 to -8.8

Note. Numbers of firms in each subsample are given in parentheses.
Source: Authors' calculations

ductions in the relative price of labor, which reduced unit costs 6.8 percent (t-ratio 2.39); and productivity gains, which reduced unit costs 8.5 percent (t-ratio 1.27).[14] The same patterns emerge from the medians and weighted averages; hence our results are robust with respect to measure of central tendency.[15]

14. Several other studies of productivity growth among Cameroonian manufacturers have been based on sector-level price deflators and have found smaller average rates of productivity growth (Biggs and Srivastava 1996, Bigsten and others 2000) The one study that uses the same firm-level deflators we use here (Tybout and others 1997) arrives at the same figure of 8.5 percent.

15. Qualitatively, the patterns match almost exactly. The only exception is the intermediate input price effect, which does not show up in our weighted averages

TABLE 8. Traditional Effective Rates of Protection: Unweighted Averages

	All	Food	Textile	Wood product	Metal & metal product	Non-exporters	Exporters	Domestic input-intensive	Imported input-intensive	Small	Medium	Large
Number of firms	34	14	9	4	7	24	10	18	16	16	11	7
Only imports tradable												
ERP 1992/93												
Mean	1.60	1.46	1.83	2.27	1.18	1.58	1.65	1.73	1.45	1.84	1.60	1.04
SD	0.90	0.94	1.01	0.50	0.64	0.87	1.02	0 70	1.08	0.79	1.06	0.69
Maximum	4.02	3.20	4.02	2.88	2.21	3.20	4.02	3.20	4.02	3.20	4.02	2.45
Minimum	0.22	0.22	0.70	1.79	0.45	0.22	0.70	0.70	0.22	0.22	0 52	0.30
ERP 1994/95												
Mean	0.59	0.68	0.51	0.80	0.42	0.69	0.37	0.73	0.45	0.78	0.40	0.46
SD	0.34	0.40	0.26	0.17	0.27	0.34	0.19	0.33	0.28	0.36	0.26	0.08
Maximum	1.69	1.69	0.94	1.01	0.79	1.69	0.66	1.69	1.20	1.69	0.96	0.59
Minimum	0.10	0.23	0.10	0.63	0.11	0.23	0.10	0.31	0.10	0.23	0.10	0.37
All inputs tradable												
ERP 1992/93												
Mean	1.15	0.92	1.62	1.09	1.02	1.01	1.48	0.96	1.35	1.08	1.34	0.98
SD	0.81	0.67	1.14	0.25	0 66	0.67	1.05	0.46	1.07	0.68	1.04	0.75
Maximum	4.02	2.33	4.02	1.46	2.21	2.33	4.02	2.29	4.02	2.29	4.02	2.54
Minimum	0.24	0.24	0 49	0.92	0.45	0.24	0.58	0.30	0.24	0.24	0.41	0.26
ERP 1994/95												
Mean	0.35	0.37	0.36	0.29	0.34	0.39	0.25	0.33	0.37	0.40	0.26	0.37
SD	0.17	0.15	0.18	0.03	0.25	0.14	0.20	0.14	0.20	0.16	0.16	0.19
Maximum	0.79	0.68	0.61	0.32	0.79	0.79	0.61	0 68	0.79	0.79	0.46	0.61
Minimum	0.03	0.03	0.10	0.25	0.07	0.23	0.03	0.03	0.07	0.23	0.07	0.03

Source: Authors' calculations.

What explains the signs and magnitudes of these effects? The tax reforms reduced unit costs because, as noted in connection with table 1, prereform turnover taxes were a larger fraction of total sales than postreform value-added taxes. Nonetheless, the impact of the domestic tax reforms was small because most of the prereform fiscal privileges took the form of tariff reductions. The significant reduction in relative labor costs is also unsurprising because nominal wages typically take some time to adjust to major devaluations. However, it is remarkable that productivity tended to improve rather than decline given the magnitude of the reduction in effective protection and the associated profit margin squeeze.

The large effect of the commercial policy reforms reflects a drop in the average nominal tariff rate on outputs from 68 percent to 27 percent, combined with a much smaller drop in the average nominal tariff rate on inputs (from 21 percent to 17 percent). The reforms had a greater effect on tariffs on products that the firms sold than on products they bought because protection levels on imported intermediate goods were already relatively modest before the reforms. This liberalization effect is also apparent in table 8, which shows that the effective protection measures fell 80–100 percentage points on average, depending on whether all inputs and outputs were treated as perfectly tradable (upper panel) or nonimported inputs were treated as nontradable (lower panel).[16]

Table 8 also reveals that the amount of cross-firm dispersion in effective protection dropped dramatically with the reforms. The cross-firm standard deviation in effective protection rates was a whopping 0.90 before the reforms, and the firm-specific values ranged from 0.22 to 4.02. After the reforms the standard deviation was 0.34 and the range was 0.10–1.69. This leveling of the playing field was due largely to the elimination of special exemptions.

One issue that often arises in Africa is whether policy reforms tend to work at cross-purposes. This appears to have occurred to some extent in Cameroon. The removal of implicit subsidies that took place with the commercial policy reforms was somewhat offset by the domestic tax reforms and the exchange rate devaluation. Nonetheless, the reductions in effective protection and the devaluation did systematically change the returns to tradable versus nontradable goods production.

V. Disaggregated Findings

We now explore the contributions of various subgroups of firms to the sample-wide summary statistics discussed. Our objective is to determine whether particular types of producers systematically did relatively well or poorly.

16. These figures are not "traditional" in the sense that firm-specific input shares are used rather than an economy-wide input-output table. In keeping with convention, these figures describe the percentage change in value added (rather than the percentage change in cost per unit revenue), so neither set of calculations is directly comparable to the fourth column of table 4 Specifically, the percentage change in costs due to tariff reforms has been divided by the share of value added in gross output

Exporters versus Nonexporters

Breakdowns by market orientation reveal that, on average, firms that were exporting in 1992/93 resembled nonexporters in most respects, although they avoided the increases in relative intermediate input prices that nonexporters suffered. This contrast probably reflected the 100 percent CFA devaluation against the French franc that took place between the sample years, which should have driven up their output prices relatively rapidly. It probably also reflected some general equilibrium effects due to the commercial policy reforms.

If we divide our sample of firms according to whether they exported in 1994/95, the contrast is more dramatic. Exporters in this group did not do unusually well in terms of their relative input prices. Nonetheless, they avoided unit cost increases altogether, mainly because they managed to increase their productivity by 19.1 percent (*t*-ratio 2.17). (They also registered rapid output growth, on average, although it was not statistically significant.) Qualitatively, the same picture emerges from output-weighted averages and medians. The fact that firms exporting at the end of the sample period did better than firms exporting at the beginning of the sample period probably reflects self-selection effects. Firms that experienced cost reductions tended to begin exporting, and those that experienced cost increases tended to cease (Clerides and others 1998). Another interpretation is that the prereform incentive structure induced some firms to export products that were not to the country's comparative advantage.

Imported Input–Intensive Firms

Producers who relied relatively heavily on imported inputs fared a bit better than those that did not, but the contrast was not statistically significant. Several opposing forces were at work. First, as one might expect, the import-intensive group was hurt a bit less by the commercial policy reforms.[17] Second, and also as one would expect, the devaluation raised their intermediate input prices, whereas these prices remained stable relative to output prices for the firms that sourced their inputs domestically. Finally, labor costs relative to output prices fell relatively rapidly among the import-intensive producers.

The same pattern emerges from sample medians (table 6), but output-weighted figures tell a somewhat different story about the relative performances of the two sets of firms (table 5). The output-weighted figures show larger cost increases for import-intensive producers and smaller cost increases for domestic input–intensive producers. However, the contrast does not stem from direct commercial policy or fiscal policy effects. It is due to pretariff intermediate input prices, which are sensitive to whether averages are weighted, because large import-intensive firms experienced major adverse shocks.

17. One reason why we do not record larger disparities is that the net tariff effect presumes perfect arbitrage between domestic and imported inputs. Hence, regardless of whether firms actually imported their inputs, they are assumed to benefit equally from liberalization-induced price reductions.

Sector-Based Breakdowns

On average wood-sector firms recorded greater unit cost growth than other firms, with a 23 percent increase. This finding stems from a large commercial policy–based reduction in output prices, which more than offset the relatively large improvements in pretariff relative prices and relatively small domestic tax effects.[18] Output-weighted averages and median figures reveal that larger wood-sector firms fared better than smaller ones, making sectorwide output growth positive.

In the food sector, weighted averages and median figures indicate that larger firms experienced substantial increases in unit costs. Despite a smaller commercial policy impact on the food sector and larger productivity gains than in other sectors, this sector was hit more severely by pretariff intermediate input price effects, which increased unit costs by 16 percent (output-weighted). Textiles and metal products experienced relatively modest cost increases, despite substantial reductions in protection, partly because they realized large productivity gains (see table 7).

VI. Summary and Conclusions

We quantified several basic changes in the incentive structure that resulted when a maxi-devaluation was accompanied by substantial tariff reductions and a major simplification of the tax structure. First, the combined effect of these changes was to increase average costs per unit revenue by 8 percent. Second, the main force driving up unit costs was the commercial policy reform, which reduced nominal protection rates on outputs much more rapidly than protection rates on inputs. The cross-firm dispersion in effective protection rates also fell markedly. Thus, despite the presence of other shocks, Cameroon's trade reforms appear to have created clear new signals for manufacturers. Third, tax reforms, reductions in the relative price of labor, productivity growth, and changes in the domestic tax structure cushioned the effects of the trade liberalization on profit margins. Finally, the CFA devaluation shifted relative prices in favor of exportable goods. Hence, as with commercial policy, the new exchange rate regime shifted the incentive structure as intended at the ground level.

Firm-level panel data allowed us to measure the effects of the policy reforms on different types of firms with considerably more precision and detail than aggregate data afford. We hope that this study provides a useful methodological example for researchers and policymakers concerned with the consequences of similar reform packages.

18. In our sample in 1994/95, 6 of the 15 food producers, 6 of the 9 textiles producers, and 4 of the 9 metal producers are exporters. However, none of the four wood producers in the sample is an exporter. In this regard, our sample is not representative of the full RPED population Indeed, despite the fact that 7 of the 40 wood sector firms in the RPED survey are exporters, these firms did not give us complete price and tax information and hence do not appear in the resurveyed sample. Our sample of wood producers is thus composed of nonexporters who did not experience much increase in their output price.

Appendix. Fiscal and Commercial Policy in Cameroon before and after Reform

In this appendix we provide further details on the fiscal and commercial policy regimes that prevailed before and after the reforms of 1994.

The Fiscal Environment before 1994

Before 1994 firms that did not enjoy access to any of the special programs offered by the UDEAC or Cameroon were subject to a variety of direct and indirect (sales) taxes (see World Bank 1992, 1995 and Gauthier and Gersovitz 1997). They included:

- *Impôt sur le chiffre d'affaires intérieur* (ICAI): Businesses in all UDEAC countries were subject to a domestic sales tax. The ICAI in Cameroon was generally levied at a rate of 10.9 percent on sales value but a reduced rate of 4.5 percent and a special rate of 2.5 percent for bakeries also existed. The ICAI was a cascading tax, because it was imposed on the value of the good or service at each level of the production process and not only on the value added. This cascading effect meant that the production tax increased with the number of intermediaries.
- *Impôt sur les bénéfices industriels et commerciaux* (BIC) and *impôt minimum forfaitaire* (IMF): Cameroon businesses were also required to pay a company tax based on the highest of the following taxes: the BIC, a tax on profits imposed at a rate of 38.5 percent (including a 10 percent communal tax) for incorporated businesses and 24.2 percent for unincorporated businesses; a 1 percent tax on sales; and the IMF, a minimum presumptive tax. Businesses in their first two years of operation were exempt from the IMF (600,000 CFA franc) and the 1 percent tax, but they had to pay the franc. During years three and four, they paid half the IMF and 1 percent tax.
- *Contribution des patentes:* Cameroon businesses were required to pay a *patente*, a kind of business license fee collected annually to help finance local governments. This tax was based on broad business activity indicators (output, equipment, number of employees).
- *Impôt spécial sur les sociétés:* Cameroon corporations were subject to a special tax, applied to capital at rates of 0.5–1.5 percent. A variety of other registration fees and taxes was also applicable. These included a registration fee for corporate charters charged at a rate of 0.25–2.0 percent, according to the firm's level of capital, and a proportional tax on income from securities (*taxe proportionnelle sur les revenus de capitaux mobiliers*) for corporations paying dividends or fees to associates and shareholders. Residents faced a tax rate on their dividend income of 16.5 percent, and nonresidents faced a rate of 25 percent. Other taxes included duties on property leases, labor housing rental, stamp duties, advertising fees, and tax licenses on land, mining, and forests.

- Taxes on insurance contracts, a trade union income tax, an apprenticeship tax, and various community taxes.

Producers subject to full taxation who engaged in international trade faced additional fiscal obligations. Imports of intermediate goods were subject to four taxes, the first three dictated by UDEAC norms and the fourth created by Cameroon. The *droits de douanes* (DD) was applied at rates of 5–30 percent on all products, regardless of origin. The *droits d'entrée* (DE) applied to all products and origins, at rates of 5–90 percent, although certain goods were exempt. The *taxe sur le chiffre d'affaires à l'importation* was imposed at a rate of 10 percent of the CFA franc value plus DD + DE. The *taxe complémentaire à l'importation* was charged ad valorem, at rates of 0–100 percent. Imports were also subject to other taxes, including an unloading fee, a municipal tax, a tax imposed by the Conseil National des Chargeurs, a tax on meat inspection, a veterinary tax, and a special tax on fuel.

Special treatment from the tax authorities was available to manufacturers through a variety of mechanisms, on a case-by-case basis. These included:

- *Taxe unique* (TU): Originally designed as a means of encouraging industrialization and trade between UDEAC countries, the TU offered firms several advantages. Qualifying firms were exempt from the domestic sales tax (ICAI), which was replaced by a firm-specific TU rate. The TU rate also replaced the tariff system. Furthermore, the TU granted preferential access to export markets in other UDEAC countries because products were exempt from duties. Neither the ICAI nor the TU tax were collected on sales to other firms with TU status. TU rates were negotiated on a firm-specific basis, and different firms may thus have paid different rates for the same product. In addition, the same firm would pay different rates on its products, depending on the country to which they were exported. To obtain TU status, firms applied to the Management Committee of the UDEAC Secretariat.
- *Taxe intérieure à la production* (TIP): Access to the TU proved difficult, so Cameroon created a domestically administered variant. This special regime also provided sales tax and tariff advantages, but in contrast to the TU, it did not give preferential access to the UDEAC market. Benefits and rates were negotiated with the Cameroon Ministry of Finance instead of with the UDEAC.
- *Investment Code:* Major tax concessions were also available under the Investment Code. Augmented in 1990 with the help of the Foreign Investment Advisory Service of the International Finance Corporation and the U.S. Agency for International Development, the Investment Code provided tax exemptions and reductions for firms meeting certain criteria. Five different schemes existed: the basic regime, the small and medium-sized enterprise regime, the strategic enterprise regime, the reinvestment regime, and the free trade zone regime. (For more details on the eligibility criteria and the benefits associated with each regime, see Gauthier and others 1995.) In

contrast to the TU/TIP rates, which could be negotiated with the authorities, Investment Code benefits were supposedly nonnegotiable. However, benefits under the TU/TIP and Investment Code regimes were not mutually exclusive, so a firm could benefit under more than one scheme at once.

- *Zone franc and point franc:* Free trade zones were part of the Investment Code in 1990 but were covered by separate legislation and administered by a separate organization. To be eligible for a free trade zone, a firm had to export 80 percent of its output and its activities had to be eligible for the basic Investment Code regime. The firm had to be located in an industrial free zone or be designated *point franc industriel* (factory-specific free zone) if it needed to be adjacent to raw materials. Free trade status brought full exemption from international and indirect taxes, and profit taxes were imposed at a reduced rate.
- *Convention spéciale:* Firms that did not find special tax schemes suited to their own specific needs could negotiate directly with the Ministry of Finance to establish a *convention spéciale* (special agreement). No guidelines existed regarding the benefits and exemptions available under such agreements, and in theory a firm could have obtained full exemption from all tax obligations, including the *patente*, for its lifetime. This unusual tax scheme was generally reserved for public or very large enterprises.

The Fiscal Environment after 1994

On January 24, 1994, Cameroon issued decrees implementing fiscal and trade reforms. The reforms included four components affecting external trade:

- *Tarif extérieur commun* (TEC): The four types of tariffs were replaced by a unified single system known as the TEC, applicable to imports from non-UDEAC countries. All external trade privileges under the Investment Code and special production regimes (TU, TIP, *conventions d'établissement*) were eliminated.
- *Reduction of tariff rates:* Imports were classified into four categories, with tariff rates ranging from 5 to 30 percent, down from 0–500 percent under the previous system.
- *Tarif préférentiel généralisé* (TPG): A general preferential tariff was introduced for trade with UDEAC countries with an initial rate fixed at 20 percent of the applicable TEC. This rate was to be reduced to 10 percent on January 1, 1996, and eliminated altogether on January 1, 1998. (The phase-out was indeed implemented.)
- A mechanism was created for charging a temporary surtax of not more than 30 percent on a set of products previously covered by quantitative restrictions and a list of designated products.

With respect to indirect taxes, the reform essentially replaced the various sales taxes with a value-added tax and eliminated special privileges. These measures included the following:

- Elimination of all indirect tax privileges under the special production regimes (TU, TIP, *conventions d'établissement*) and the Investment Code, except the free trade zone.
- Introduction of a *taxe sur le chiffre d'affaires* (TCA), a quasi–value-added tax applicable to domestic production and to imported inputs and intermediates, replacing the former sales and production tax (ICAI, TU, TIP). (We use the term "quasi" because firms initially paid taxes on their purchases, then periodically applied to the government for reimbursement.) Three categories of products were specified: those subject to the normal rate (12.5 percent, increasing to 15 percent on January 1, 1995, and to 17 percent in 1996); those subject to the reduced rate (5 percent, increasing to 8 percent on January 1, 1995); and exempted goods.
- Creation of a mechanism for applying excise taxes to certain products.

On February 1, 1994, the reform went into force for firms governed by the common law system. Firms receiving special fiscal privileges were allowed a transition period. Those governed by the IC, TU, and TIP were not subject to the new regime until the 1994/95 fiscal year (beginning July 1, 1994). Firms governed by special agreements were given until December 31, 1995, to regularize their situation. This period of negotiation was later extended to March 31, 1996.

The reforms left the free trade zone intact. Hence qualifying firms continue to enjoy full exemption from import duties and TCA, and they are exempt from income taxes in the first 10 years of their existence. Firms that already existed before the free trade zone was created pay an income tax of 15 percent instead of the normal rate (38.5 percent). Exporters not in the free trade zone can apply for refunds of a portion of the customs they pay on imported inputs. The fraction refundable is equal to the share of their total sales exported outside the UDEAC. However, given the inefficiency of the administration and thedelays in paying tax credits, this benefit has proved of little use to marginal exporters.

References

Balassa, Bela. 1965. "Tariff Protection in Industrial Countries: An Evaluation." *Journal of Political Economy* 73(6):573–94.

Biggs, Tyler, and Pradeep Srivastava. 1996. "Structural Aspects of Manufacturing in Sub-Saharan Africa: Findings from a Seven-Country Enterprise Survey." World Bank Discussion Paper 346. Washington, D.C.

Bigsten, Arne, Paul Collier, Stefan Dercon, Marcel Fafchamps, Bernard Gauthier, Jan Willem Gunning, Jean Habarurema, Abena Oduro, Remco Oostendorp, Catherine Pattillo, Måns Soderbom, Francis Teal, and Albert Zeufack. 2000. "Exports and Firm-Level Efficiency in African Manufacturing." Working Paper Series 2000-16. University of Oxford, Center for the Study of African Economies, Oxford, UK.

Clerides, Sofronis, Saul Lach, and James Tybout. 1998. "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico and Morocco." *Quarterly Journal of Economics* 113(3):903–47.

Gauthier, Bernard, ed. 1995. *Economic Reform and the Development of the Manufacturing Sector in Cameroon.* RPED Paper 44. World Bank, Washington, D.C.

Gauthier, Bernard, and Mark Gersovitz. 1997. "Revenue Erosion through Exemption and Evasion in Cameroon, 1993." *Journal of Public Economics* 64(3):407–24.

Gauthier, Bernard, Jean-Pierre Legoff, and Luc Vallée. 1995. "State Crisis and the Manufacturing Sector in Cameroon." Monographs on International Business and Economics, No 95-04, CETAI, HEC, Montreal.

Head, Keith, and John Reis. 1999. "Rationalization Effects of Tariff Reductions." *Journal of International Economics* 47(2):295–320.

Michaely, Michael, Demetris Papageorgiou, and Armeane Choksi. 1991. *Liberalizing Foreign Trade: Lessons of Experience in the Developing World.* Vol. 7. Cambridge, Mass.: Basil Blackwell.

Tybout, James, Bernard Gauthier, Giorgio Barba Navaretti, and Jaime de Melo. 1997. "Firm-Level Responses to the CFA Devaluation in Cameroon." *Journal of African Economies* 6(1):3–34.

World Bank. 1992. *Regional Cooperation for Adjustment: A Program of Trade and Indirect Tax Policy Reforms for Member Countries of the Customs Union of Central African States* (UDEAC). Report No. 9747-AFR. Washington, D.C.

———. 1995. *Republic of Cameroon. The Challenge: Harnessing Unrealized Potential (A Private Sector Assessment).* Report No. 13955-CM. Washington, D.C.

The World Bank Economic Review
Author Index to Volume 16, 2002

Number 1: 1–148
Number 2: 149–319
Number 3: 321–472

The World Bank Economic Review
Title Index to Volume 16, 2002

Number 1: 1–148
Number 2: 149–319
Number 3: 321–472

New books from *OXFORD*

Valuing Freedoms

Sen's Capability Approach and Poverty Reduction

Sabina Alkire

352 pages March 2002
0-19-924579-7 £45 00 $70 00 Hardback

NEW IN 2003

From Conflict to Recovery in Africa

Edited by Tony Addison

UNU/WIDER Studies in Development Economics

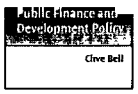
312 pages March 2003
0-19-926103-2 £50 00 $80 00 Hardback

NEW IN 2003

Public Finance and Development Policy

The Annual Bank Conference on Development Economics - Europe

Clive Bell



300 pages March 2003
0-19-877367-6 £21 99 Paperback
0-19-877366-8 £50 00 Hardback

NEW IN PAPERBACK 2003

Population Matters

Demographic Change, Economic Growth, and Poverty in the Developing World

Edited by Nancy Birdsall, Allen C. Kelley, and Steven Sinding

456 pages January 2003
0-19-924407-3 £45 00 $72 00 Hardback 2001
0-19-925631-4 £19 99 $29 95 Paperback

NEW SECOND EDITION

India

Development and Participation

Jean Drèze and Amartya Sen

540 pages August 2002
0-19-925748-5 £45 00 $72 00 Hardback
0-19-925749-3 £18 99 $24 95 Paperback

NEW IN PAPERBACK 2002

Peasants Versus City-Dwellers

Taxation and the Burden of Economic Development

Raaj K. Sah and Joseph E. Stiglitz

238 pages February 2002
0-19-925748-5 £30 00 $49 95 Hardback 1992
0-19-925749-3 £13 99 $19 95 Paperback

www oup com

Credit card Hotline
+44 1536 454534

book orders@oup com

OXFORD
UNIVERSITY PRESS

Coming in the next issue of

THE WORLD BANK ECONOMIC REVIEW

Volume 17, Number 1, 2003

A NEW DEVELOPMENT DATABASE

THE WORLD BANK
1818 H Street, NW
Washington, DC 20433, USA
World Wide Web: http://www.worldbank.org/
E-mail: wber@worldbank.org

ISBN 019-851662-2

UPC 0 78019 83386 5 24