NOTE NUMBER 352

MACROECONOMICS, TRADE & INVESTMENT GLOBAL PRACTICE MARCH 2018

THE WORLD BANK GROUP

viewpoint
PUBLIC POLICY FOR THE **PRIVATE SECTOR**

Trade Facilitation Reform

The Impact of Targeted Sampling on Import Compliance

The results of an impact evaluation of the introduction of a risk-based approach to sampling import shipments by the Macedonian Food and Veterinary Agency indicate that the sampling plan was not effectively implemented, and the reform did not improve the targeting of risky shipments. Evidence that the reform increased trade flows is inconclusive. The weak results suggest a need to concentrate efforts in improving data management in technical agencies, not only for analytical purposes but especially to improve the design and effectiveness of the reforms.

Russell Hillberry, Ana Margarida Fernandes, and Alejandra Mendoza Alcántara

Russell Hillberry (rhillber@purdue.edu) is an Associate Professor in the Department of Agricultural Economics at Purdue University. Ana Margarida Fernandes (afernandes@worldbank.org) is a Senior Economist in the Trade and International Integration Unit of the Development Research Group at the World Bank. Alejandra Mendoza Alcántara (amendoza1@worldbank.org) is an Impact Evaluation Specialist in the Finance, Competitiveness and Innovation Global Practice at the World Bank Group.

Research for this note was supported by the governments of Canada, the United Kingdom, and the United States through the Impact Program of the World Bank Group's Trade and Competitiveness Global Practice. Additional support was provided by the World Bank's Multi-Donor Trust Fund for Trade and Development, the Strategic Research

(continued on back page)

WORLD BANK GROUP

Every year hundreds of millions of dollars of development aid are spent on trade facilitation projects. Few projects are subjected to scrutiny via formal impact evaluations, and most of the studies that have been conducted evaluate reforms by customs agencies.[1]

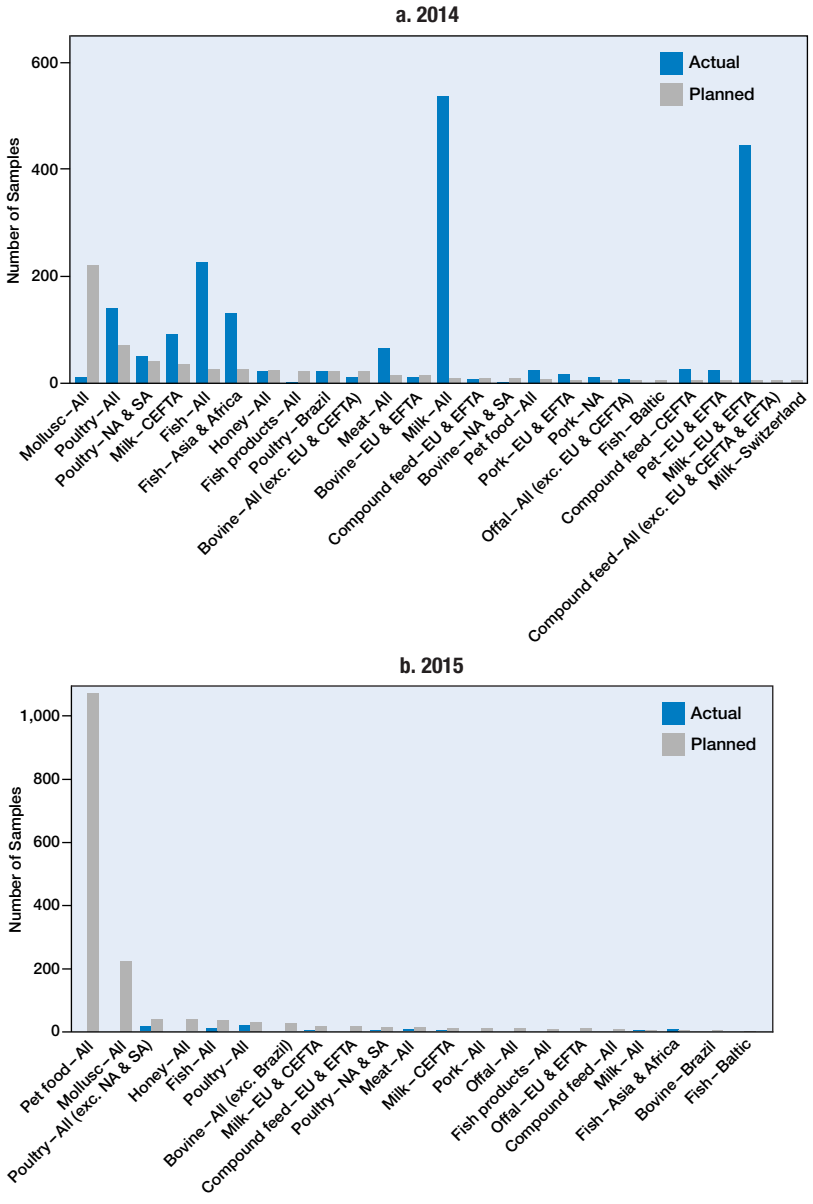
This note is part of a research series on trade facilitation reforms in the Western Balkans. It is the first study to evaluate a trade facilitation reform undertaken by a "technical agency" (an agency tasked with oversight of food safety, environmental protection, or other responsibilities beyond the remit of the customs agency). The study evaluated the introduction of a risk-based approach to sampling of import shipments for the purpose of laboratory testing (Fernandes, Hillberry, and Mendoza 2017). The reform was undertaken in 2014 by the Food and Veterinary Authority (FVA) of the former Yugoslav Republic of Macedonia (FYR Macedonia).

Prior to the reform, the World Bank Group's Western Balkans Trade Facilitation project identified lengthy and intrusive sampling and laboratory testing regimes by technical agencies as a significant constraint on trade in FYR Macedonia. Laboratory tests of sampled shipments very rarely identified noncompliant shipments, suggesting poor targeting and excessive sampling. In 2013, for example, samples were taken from nearly 1 in 10 import shipments. Less than 0.2 percent of the samples taken were found to be noncompliant (authors' calculations based on FVA administrative data for imports of food of animal origin and animal feeds). High rates of sampling subjected imports to queuing at the border, causing delays during the movement of goods. Shippers of goods that were sampled for testing also had

to wait for laboratory testing results before their goods could be released to the market.

The research reviewed here used data on planned and actual sampling activity, laboratory test outcomes, and international trade flows to assess the impact of the reform. It provides answers to three questions: How closely did actual sampling activity match the sampling activity planned in the reform? Did the reform cause sampling activity to become better linked to the probability that samples do not comply with the FVA's technical standards? How did the reform affect import flows?

Figure 1 **Planned and actual number of import shipments sampled by the Food and Veterinary Authority, by product and country group, 2014 and 2015**



Source: Fernandes, Hillberry, and Alcantara 2017.
Note: Figures show information based on sampling plans; the FVA laboratory samples dataset; and the single window for import, export, and transit of goods and tariff quota (EXIM) dataset. The datasets are combined based on the broad product group and cover food of animal origin and animal feeds.
Group of countries: All countries (All), North America (NA), South America (SA), European Union (EU), Central European Free Trade Agreement members (CEFTA), European Free Trade Association members (EFTA).

Improving sampling procedures in the FVA

The goal of the reform was to reduce administrative and private sector costs through adoption of a risk-based approach to sampling import shipments for laboratory testing. With World Bank Group support, in 2013 FVA staff received training on risk-based approaches. Training included the development of risk scores and guidance on constructing an annual sampling (or monitoring) plan based on historical data. The objective of this process was to enable the FVA to better focus sampling and testing efforts on shipments that were more likely to be noncompliant with technical standards and more dangerous if noncompliant. More focused sampling activity would improve oversight of import shipments while reducing the number of shipments sent for laboratory testing. The sampling plan outlined the number of samples to be taken the following year for a given border post, product group, and source country or group of countries. Sampling plans were designed for three sets of products under FVA oversight: food of animal origin, animal feeds, and food of nonanimal origin and food contact materials.

Weak adherence to sampling plan

The annual sampling plans that were designed as part of the reform aim to optimize sampling activity by border agents, taking into account expected probabilities of noncompliance and potential degrees of harm, given resource constraints. Each annual plan is intended to guide sampling activity over the subsequent year, although agents can deviate from the sampling plan as necessary. The discovery of salmonella in products imported into another country in the European Union, for example, would likely generate additional FVA sampling activity on similar shipments. Even absent news of a bacteria outbreak, FVA border agents retain discretion to depart from the sampling plan when they suspect that a shipment warrants additional scrutiny.

Raw summary statistics reveal sizable differences between actual sampling activity and the sampling plans. The sampling plan for 2014 projected that 598 samples would be taken from two classes of products: food of animal origin and animal feeds. The actual number of samples taken surpassed the plan by a factor of three (figure 1a). In contrast, in 2015 the actual number of samples taken was 23 times smaller than the planned number (figure 1b).

A regression model was used to assess the correlation between the actual samples taken in 2014 and 2015 and the corresponding sampling plans.[2] The analysis was limited to the two sets of products subject to FVA oversight for which complete data are available: food of animal origin and animal feeds. The regression controlled for the inherent risk characteristics associated with each group of products, group of source countries, and year.[3]

The results indicate weak adherence to the sampling plan by FVA border agents. A positive coefficient found on the planned number of samples suggests that deviations from the sampling plan leaned in the direction of oversampling shipments that had initially been deemed most risky (those with higher numbers of planned samples).

Two unexpected challenges to implementation of the reform plan help explain these results. Oversampling in 2014 can be attributed to the fact that implementation of the reform was delayed until July of 2014. Actual sampling rates in 2014 therefore included six months of sampling activity that preceded the plan. In 2015, a funding dispute between the FVA and the laboratories that test the samples dramatically limited sampling activity.

FVA failed to target risky shipments

One of the objectives of the risk-based sampling reform was to align sampling activity with the likelihood that shipments would be judged noncompliant by subsequent laboratory tests. If the reform was effective, a stronger positive relationship between noncompliance rates and sampling rates would be observed after the reform than before it. A regression model was used to test this effect.[4] The estimation controlled for unrelated factors, such as differences in product characteristics or the potential for harm.[5]

A preliminary analysis showed that the number of shipments that did not comply with the FVA's technical standards in 2013–15 was extremely small (fewer than 40). In the second semester of 2014, sampling and noncompliance rates increased for food of animal origin, possibly because of the 2013 aflatoxin outbreak (table 1).

Given the very small number of noncompliant

Table 1 **Sampling and noncompliance rates of all imports, 2013–15**

Year	Number of import shipments sampled	Total number of import shipments	Number of sampled import shipments noncompliant	Sampling rate (percent)	Non-compliance rate (percent)
2013					
First half	835	8,167	12	10.3	1.4
Second half	446	8,479	1	5.3	0.2
Total	1,281	16,646	13	8	1
2014					
First half	451	7,932	2	5.7	0.4
Second half	780	8,643	21	9.1	2.7
Total	1,231	16,575	23	7	2
2015					
First half	4	8,844	0	0	0
Second half	59	9,615	0	0.6	0
Total	63	18,459	0	0	0

Source: Fernandes, Hillberry, and Alcantara 2017.
Note: Table shows information based on the FVA laboratory samples dataset and the EXIM dataset. The datasets are combined using the broadly defined product grouping and cover food of animal origin and animal feeds.

shipments, the regression estimates were unable to identify a significant relationship between the likelihood of noncompliance and the sampling rate before the 2014 reform. Considering also data from the post-reform period, the analysis reveals that the probability of a shipment being noncompliant was lower when sampling rates were higher, especially after the reform.

This result is concerning, because risk-based targeting aims to focus sampling activities on shipments that are more likely to be noncompliant. One would therefore expect the probability of noncompliance among sampled shipments to be higher when sampling rates are higher.[6] The findings suggest that the targeting of sampling activity associated with imports of food of animal origin and animal feeds was not effective in FYR Macedonia.

A caveat with these and other exercises on sampling activity is that the data needed to calculate sampling rates for less risky products (food of nonanimal origin and food contact materials) were not collected. Given that those products are less likely to pose health risks to the public if noncompliant, the FVA reform was designed to reduce their sampling rates by larger magnitudes. The reform may have succeeded in ways that could not be measured directly.

Limited impact of reform on trade

A central question typically asked in the small literature evaluating trade facilitation reforms is the degree to which such reforms affect trade. In this study, this question was addressed in two ways. First, an exercise examined whether imports of products subject to FVA oversight saw more rapid growth than other products following reform. Second, an exercise tested whether larger reductions in the sampling rate led to larger increases in the import value of products subject to FVA oversight.

In the first exercise, a difference-in-difference regression was used to compare import growth of products subject to FVA oversight with growth of similar products. It found that the import value of FVA products rose 3–5 percent following the reform. These estimates include the effects of any other FVA policy changes implemented at the same time as the reform. They may also be conflated with unobservable changes in the market for FVA products that coincided with the reform (though efforts were undertaken to control for export supply and import demand shocks and to restrict the sample of control products to be more "similar" to the FVA products to rule out this possibility). If the estimates of an additional 3–5 percent growth in imports of FVA products are correct, trade response parameters suggest that the reform had an impact roughly equivalent to a 1-percentage point tariff cut on FVA products.

The second exercise used a regression model to assess the relationship between changes in the sampling rate and changes in import value. This exercise was complicated by the 2015 funding dispute with the laboratories, which caused a dramatic drop in the sampling rate that is not attributed to the reform. The relationship between sampling rates and import value is of interest to the trade facilitation literature, so it was estimated with and without the 2015 data, which include the effects of the funding dispute. The exercise found no evidence of an effect of reduced sampling on import value over the 12–18 months immediately after the reform.

Overall the estimates indicate inconclusive effects of reduced sampling rates on import value. Changes in the sampling rate had no immediate effect on import value; it is possible that such effects emerge with a lag over longer time horizons. The findings reveal that imports of FVA products grew slightly faster than various sets of control products following the reform, though it is not possible to say conclusively that this growth was a result of the reform. These findings need to be interpreted in light of the caveat that it was not possible to estimate relationships between sampling rates and imports for less risky products, for which larger changes in the sampling rates were planned.

Technical agencies need to improve data collection efforts

This research highlights the challenges associated with an impact evaluation of reforms undertaken by technical agencies that play an important role at the border but are unable to collect comprehensive data on variables of interest. One reason data are lacking is that technical agencies sometimes collect data in ways that are not consistent with the international trade

classification system. A lack of comprehensive data on test outcomes also limited the evaluation of substantial reforms that occurred in the sets of products imposing smaller public health risks. In the FVA, these data challenges not only undermine evaluation efforts, they also limit the effectiveness of both the risk-based system and the construction of the sampling plans.

Recommendations for improving data management include:

1. The FVA should itself collect all test outcomes—positive and negative—in an electronic format or platform, rather than relying on the laboratories to collect and provide this information. The information collected from each import shipment sampled should be linked to the shipment identification number issued by the single window for import, export, and transit of goods and tariff quota (EXIM), which collects information on the volume of imported products, the country of origin, and the firm engaged in each import transaction. Electronic data collection would allow more precise tracking of the characteristics of noncompliant shipments and the updating of the sampling plan in real time.

2. A more ambitious data management reform would seek to integrate information from laboratory test results with data recorded by other agencies, especially the customs agency. Data at the shipment level could be used to improve the targeting strategy and the sampling plan. Incorporating shipment-level data contained in the customs declaration would require high levels of coordination with the national customs agency.

3. Sampling plans should be updated more than once a year and take into account emerging reports of outbreaks or evidence of failed outcomes at other agencies. Emerging risks are currently communicated to border agents through other channels; updating the sampling plan accordingly makes sense. Frequent updating of the sampling plan would better tie actual sampling behavior to the plan and help the agency learn about the effectiveness of its approaches to updating the plan.

Conclusion

This study is innovative for three reasons: (a) it applies rigorous impact evaluation methods to trade facilitation reforms; b) it explores the impact of the reform on compliance with import regulations, an outcome the literature previously had ignored; and c) it assesses for the first time a reform implemented by a technical agency. The study highlights the need to concentrate efforts in improving data management in such agencies, not only for analytical purposes but particularly to improve the effectiveness and design of subsequent reforms. In this case, a data management reform would help track the effectiveness of the sampling plan, monitor its implementation, and improve targeting based on shipment characteristics.

Many of the study's findings are inconclusive, because of the short sample period; the weak quality of the data; and specific implementation issues, such as the funding dispute with the public laboratories in 2015. But the research highlights the fact that several underlying mechanisms—including the design of the sampling plan, the testing process, and training for border agents—make trade facilitation reforms work. The contributions of these subcomponents represent knowledge gaps that could be filled by a different evaluation design.

(See back page for Notes and References.)

Notes

1. See, for example, Fernandes, Hillberry, and Alcantara (2015); Fernandes, Hillberry, and Berg (2016); Carballo, Schaur, Graziano, and Martincus (2016); and Carballo, Schaur, and Martincus (2016).
2. The regression model estimated by ordinary least squares (OLS) explains the relationship between the number of import shipments sampled and the number of samples to be taken according to the sampling plan, at the product-group of country-year level.
3. This control was handled by including fixed effects for the product group, the country group, and the year as regressors in some of the OLS regressions. The assessment also excluded outliers, such as milk and milk products, for which Europe set up a warning in 2013 following the outbreak of aflatoxin, and pet food products, for which the number of shipments to be sampled increased dramatically in 2015.
4. The regression model estimates a relationship between an indicator for noncompliance of an import shipment and the associated sampling rate. Sampling rates are calculated as the ratio between the number of shipments sampled and the total number of imported shipments for each broad product group and semester.
5. This control was handled by including in the OLS regressions fixed effects for the product group, the country group, and the semester-year.
6. The sampling and noncompliance rates were negatively correlated even in the second semester of 2014, when the funding dispute with the laboratories, which undermined the analysis of 2015, was not a factor.

References

Carballo, Jerónimo, Georg Schaur, Alejandro Graziano, and Christian Volpe Martincus. 2016. "Transit Trade." IDB Working Paper IDB-WP-704, Inter-American Development Bank, Washington, DC.

Carballo, Jerónimo, Georg Schaur, and Christian Volpe Martincus. 2016. "Trust No One? Security and International Trade." IDB Working Paper IDB-WP-703, Inter-American Development Bank, Washington, DC.

Fernandes, Ana M., Russell H. Hillberry, and Alejandra Mendoza Alcantara. 2015. "Trade Effects of Customs Reform: Evidence from Albania." Policy Research Working Paper 7210, World Bank, Washington, DC.

———. 2017. "An Evaluation of Border Management Reforms in a Technical Agency." Policy Research Working Paper 8208, World Bank, Washington, DC.

Fernandes, Ana M., Russell H. Hillberry, and Claudia N. Berg. 2016. "Expediting Trade: Impact Evaluation of an In-House Clearance Program." Policy Research Working Paper 7708, World Bank, Washington, DC.

viewpoint

is an open forum to encourage dissemination of public policy innovations for private sector–led and market-based solutions for development. The views published are those of the authors and should not be attributed to the World Bank or any other affiliated organizations. Nor do any of the conclusions represent official policy of the World Bank or of its Executive Directors or the countries they represent.

To order additional copies
contact Jenny Datoo,
managing editor,
Room F 5P-504,
The World Bank,
1818 H Street, NW,
Washington, DC 20433.

Telephone:
001 202 473 6649
Email:
jdatoo@worldbank.org

Produced by Carol Siegel

Printed on recycled paper

(continued from front page)

Program on Economic Development, and the Impact Evaluation for Development Impact (i2i) Fund.