Expediting Trade: Impact Evaluation of an In-House Clearance Program*

Ana M. Fernandes	Russell Hillberry	Claudia Berg
World Bank	World Bank	World Bank

September 2016

Abstract: Despite the importance of trade facilitation as an area of trade and development policy, there have been very few impact evaluations of specific trade facilitation reforms. We evaluate in-house clearance, a reform that allows qualified firms in Serbia to clear customs from within their own warehouse rather than at the customs office. The pooled synthetic control method applied here offers a novel solution to many of the empirical challenges that frustrate efforts to evaluate trade facilitation reforms. We estimate causal impacts of in-house clearance on trade outcomes for 21 firms that adopted it for import shipments. Our results suggest that the most evident benefit of the program for participating firms is reduced uncertainty about clearance times.

Keywords: Trade facilitation, Customs, Trade costs, Synthetic controls, Program evaluation.

JEL Classification codes: F13, F14, F15, C22.

1. Introduction

Trade facilitation is at the center of the multilateral trade agenda. The World Trade Organization's 2013 Trade Facilitation Agreement (TFA) will guide a host of operational reforms designed to increase the speed and to reduce the cost of moving goods across national borders. Most TFA-related reforms will occur in developing countries, and many will be funded by developed country aid agencies or multilateral development agencies such as the World Bank. The hundreds of millions of dollars per year in aid-for-trade funds committed to support these reforms make trade facilitation an important area of international development policy as well as trade policy.[1] Yet unlike many other areas of development policy, these reforms have received relatively little formal scrutiny via the mechanism of impact evaluations.[2] Correcting this oversight represents an important agenda for international trade policy research.[3]

Customs reform is an area of particularly significant policy interest, but also one that has seen few formal impact evaluations.[4] In this paper we evaluate one particular type of customs reform - the introduction of an in-house clearance (IHC) program by the customs agency of the Republic of Serbia. The IHC program allows qualifying firms to clear customs on their import shipments from within their own warehouse, rather than doing so at the customs office. IHC programs are recommended in the WTO Trade Facilitation Agreement, and they are often recommended for implementation by developing country customs agencies. Our evaluation therefore offers a window into the likely consequences of these reforms firms that adopt the program.

Because in-house clearance allows firms to avoid congestion or other potential sources of delay at the customs office, it offers firms the possibility of reduced customs clearance times and reduced uncertainty about clearance times. To study these effects we investigate changes in the median clearance time and in clearance times at the 75th and 90th percentiles of the distribution.

---

[1] From 2006 to 2011, governments and donors disbursed over $1.2 billion in official development assistance to support trade facilitation efforts (WTO 2013). Commitments made since the 2013 signing of the TFA have been even more substantial.

[2] Export promotion is one sub-area of trade facilitation reforms that has seen multiple impact evaluations. Export promotion is a type of intervention that lends itself to randomization as in Atkin et al. (2014), even though most of the interventions evaluated so far such as those discussed in Volpe Martincus and Carballo (2008) or Cadot et al. (2015) did not include a randomized treatment.

[3] There is a vast literature on international trade costs (see Anderson and van Wincoop 2004) which commonly assumes that all international trade costs are directly proportional to the volume of trade flows, and calculates the welfare consequences of removing such costs. The most plausible source of international trade costs that would be roughly proportional to trade volume would be transportation costs or border-related costs such as those required to deal with customs. In that sense studies such as ours are best able to inform the question of whether proportional-to-volume costs can be effectively reduced through policy interventions.

[4] The study of the impact of customs reform in Albania by Fernandes et al. (2015) is a notable exception. Volpe Martincus et al. (2015) study risk management in Uruguayan export inspections, but do not study a particular reform episode.

Reductions in the latter two statistics indicate a compression of the clearance time distribution, and thus reduced uncertainty about the time required for customs clearance.[5] We also investigate the effect of program adoption on the inspection rate. Reductions in clearance times, uncertainty or inspection rates (along with any other operational benefits) are likely to reduce (unobserved) trade costs and thus might allow participating firms to increase their imports. We therefore investigate changes in adopting firms' monthly imports.

As with many customs reforms, evaluation of the IHC reform is complicated by a number of factors. First, randomized treatment is virtually impossible to implement given the operational demands of customs and other border agencies, so only *ex post* evaluation methods are feasible. Second, financial and other costs associated with participation in the IHC program mean that relatively few Serbian importers began to use the program during the period of our data. Moreover, the participating firms differ substantially from the typical importing firm in the size and composition of their import bundle, and in the moments of the distribution of the time needed to clear their import shipments. The underlying heterogeneity of trade outcomes across treated and untreated units present important challenges for the evaluation of many trade facilitation reforms. Third, the trade outcomes we study have complex dynamics that are quite problematic for conventional impact evaluation tools such as difference-in-differences (DiD) estimators.[6]

These challenges are daunting for any empirical methodology, but mitigated by the synthetic control method proposed by Abadie et al. (2010). We believe the method is well-suited for evaluation of many customs reforms. The method can exploit the high-quality administrative data that exists in the customs environment. In particular it can capture, and exploit, the complex dynamics observed in the outcome variables of interest. The method also offers more credible treatments of heterogeneity and selection than other *ex post* evaluation methods. A recent methodology proposed by Dube and Zipperer (2015) offers a robust approach to hypothesis testing that pools synthetic control results across multiple treated units. Among other advances, the pooled approach allows heterogeneous impacts to be estimated across treated units, even as it

---

[5] It may seem unusual to use these higher percentiles as measures of uncertainty rather than, for example, the variance or the interquartile range of the clearance time distribution. The IHC program imposes a default clearance time that mechanically raises the clearance time in the lower half of the distribution for most IHC firms. In our view reductions in clearance time at the upper end of the distribution are more valuable to the firm than reductions at the low end, so we focus on statistics that inform reductions at the upper end of the distribution.

[6] Indeed, in this paper we show that DiD is invalid for an evaluation of the impact of the IHC program because the assumption of common trends is violated.

exploits the benefits of increased statistical power associated with pooling across them.[7]

Using this recently developed pooled synthetic control method our study investigates the impact of the IHC program on several firm-level outcomes that are of interest to trade facilitation practitioners and policymakers.[8] We first ask whether firms adopting the program saw reductions in the median time required for the customs agency to clear their import shipments. While the point estimates across several robustness checks indicate that the program reduced median clearance times, the estimated effects are not statistically significant. In order to investigate the effects of the program on clearance time uncertainty we investigate changes in the upper half of the distribution of clearance times. We find strong evidence that clearance times fell at the 75th percentile of the distribution and mixed evidence that they fell at the 90th percentile. Our preferred estimates suggest that clearance times fell by 40% (42 minutes) at the 75th percentile and 26% (57 minutes) at the 90th percentile. A subsequent test for cross-firm heterogeneity leads us to accept the null hypothesis that these effects are common across IHC-adopting firms. Next, we ask whether inspection rates fell for IHC-adopting firms, but find no evidence to support this. Finally, we do not find statistically significant evidence that the program increased firm-level imports within post-reform time windows of six or nine months, the windows that are possible given data availability. We construct 95 percent confidence intervals for all five of the outcome variables we study, in order to provide estimates of reasonable bounds on the magnitudes of the true treatment effects. Although the confidence intervals for median clearance times, the 90th percentile of the clearance time distribution and log import value contain zero, the bulk of these intervals lie on the side of zero that is expected.

The modest benefits of the IHC program for importing firms that are observed here might be understood in the context of the already favorable treatment that IHC firms receive in customs prior to the implementation of the program. Customs clearance times for these firms are already low, as is appropriate for firms that are understood to be low risks for non-compliance with

---

[7] Deaton (2010) argues that the assumption of no heterogeneity of impact across treated units is an important shortcoming of many program evaluation methods, including standard approaches that exploit randomized control trials. The method we employ allows counterfactual estimates at the level of the individual unit. These estimates are then pooled to test jointly for program impact, and for heterogeneity.

[8] We estimate effects of IHC at the firm level because the composition of an IHC firm's imports across products and countries is arguably endogenous to firm-level decisions about how best to use the program once it has been adopted. While greater statistical power would be available if we were to assume a common treatment effect across products and/or countries, analysis at that level would presume no tradeoffs between clearing one set of goods or another under the program. By focusing on firm-level - rather than more disaggregated - outcomes we can be agnostic about the effects of firms' strategic choices regarding use of the program on the composition of their imports across products and source countries. Our estimates summarize both direct effects of the program and second-order effects that operate through firms' operational decisions under the program.

customs regulations. The IHC program may have other benefits that are not quantified here; after all, the adopting firms do choose to adopt, despite the compliance costs that apparently deter other firms from adopting. But these other benefits do not appear to lead to increased import value, at least over the time horizon that we are able to observe. The absence of an effect on import value may indicate that the program does not increase firm-level imports, but it may also suggest that effects on trade occur over longer durations than we are able to study in this application.

The paper is organized as follows. Section 2 describes Serbia's in-house clearance program. Section 3 describes the synthetic control methods and the methods Dube and Zipperer (2015) use to pool across treated units. Section 4 describes the data. Section 5 provides results and Section 6 concludes.

2. The In-House Clearance Program in Serbia

2.1. What Is In-House Clearance?

'Trusted trader' programs - more formally known as 'authorized economic operator' programs - are a staple of well-developed customs environments. In these programs, qualifying firms benefit from a variety of special or expedited procedures available only to those firms that meet specific standards or criteria established in the customs law. Those criteria indicate that the authorized firms are exceptionally low risks for non-compliance with import regulations. Firms that trade frequently can find trusted trader programs to be quite useful. The programs also benefit the border management authorities because they allow resources to be deployed toward consignments that are higher risks for non-compliance.

One of the procedures that is sometimes incorporated in trusted trader programs is *in-house clearance* - a procedure that allows firms to clear their import or export shipments at a suitable commercial warehouse, rather than at the location of the customs office.[9] In most cases, goods cleared under an in-house clearance program are not inspected; the shipment is cleared electronically (and remotely). Only occasionally do customs officers visit the warehouse to conduct inspections of paper documents and/or the goods themselves. In this paper we conduct the first (to our knowledge) impact evaluation of an in-house clearance procedure.[10]

---

[9] These procedures are sometimes called 'off-site inspection' procedures because the inspections, if they occur, occur in the warehouse and not at the site of the customs office.

[10] Carballo et al. (2016) conduct an *ex post* evaluation of an authorized economic operator program in Mexico aimed at facilitating licit trade in face of security concerns. The Mexican program offered preferential treatment in customs but does not appear to have

The Serbian customs agency implemented an in-house clearance program in 2011. Although the program allows in-house clearance both for imports and for exports, our available data are for imports so we focus on the operation of the IHC program for imports.[11] The IHC program for imports operates as follows: when shipments that qualify for in-house clearance arrive at a Serbian border crossing point, Serbian customs clears the shipment using ordinary 'transit' procedures that are otherwise used for goods passing through Serbia.[12] Based upon the time of clearance and the expected duration of the travel to the warehouse, a deadline for completion of the procedure is set. The driver then proceeds to the approved inland warehouse.

Once the goods arrive at the approved warehouse, the firm must notify the customs office that the goods have arrived intact, declare any incidents that have occurred in transit, and issue a certificate indicating that they have received the goods from the haulier. The means of transport (e.g. the truck) must be moved within the warehouse to the area specified by the agreement with customs. The firm then sends the import declaration electronically to the Serbian customs authority.

Barring any irregularities, the customs office enters the import declaration into the system, and the declaration is processed by the automated risk management software. In the Serbian IHC program, the software will only require an inspection if the import declaration is one of the declarations that are randomly chosen for inspection.[13] From the time of the filing of the declaration, the customs agency has exactly one half hour to determine whether an inspection of the shipment is warranted, and to notify the firm. If the indication is that the shipment is to be inspected, the means of transport remains sealed and unloaded in the specified location until the customs officers arrive to the firm's warehouse to conduct the inspection. In normal circumstances the shipment is not inspected and is cleared automatically after one half hour has passed. After clearance the firm is free to unload the shipment and to begin using the goods as desired. The firm is given another deadline to submit any remaining documentation.

---

incorporated in-house clearance. Carballo et al find that the program reduced clearance times and inspection rates for exporters and increased their export values.

[11]In most cases export clearance is likely to be less burdensome than import clearance so the effect of the program is likely to be more valuable for imports than for exports.

[12] One of the ordinary transit procedures would be to seal the truck if it is not already sealed.

[13] Serbian customs employs a risk management model that selects shipments for inspection. Some shipments are 'targeted' for selection because they have characteristics that suggest higher risks of non-compliance. Other shipments are chosen randomly for inspection. Random selection helps to verify that the risk model is working properly and deters opportunistic non-compliance for shipments that are typically identified as low risk. For more on risk management in customs see Fernandes et al. (2015).

Shipments that do not qualify for in-house clearance (either because the goods are not eligible even for qualifying firms, or because the firm is not qualified) are routed instead to a customs office for clearance. The customs office may be at the border crossing point or it may be inland, depending on the particular route that is taken. For example, there is a terminal in the capital city of Belgrade (which is inland), and most shipments bound for Belgrade would go through clearance procedures at the Belgrade terminal. Congestion or other complications at the terminal may increase clearance times or cause them to vary in ways that are problematic for the importing firm. By clearing shipments at the warehouse rather than at the terminal, IHC firms can avoid these difficulties. In principle, shorter and less variable clearance times may allow firms adopting the IHC program to increase their imports if the program reduces these implicit trade costs.

2.2. Take-Up and Use

The Serbian customs agency implemented the IHC program in 2011, with the first firm to adopt it doing so in July 2011. The program is open to any firm that meets the qualification criteria that are reported in Table 1. In the import customs data we have for the period 2010-2013, 34 importing firms used the program. The first use of the IHC program to clear imports occurred in July 2011. Our data end in December 2013, and evaluation is not possible without a period of time that follows take-up of the program. We therefore limit the sample to 21 importing firms that adopted the program before August 2013 and that import in every month of the 2010-2013 period. These firms are the early adopters, and they tend to be substantially larger than the typical Serbian importing firm, as will be documented in Section 4.[14]

Discussions with those familiar with the operation of the Serbian IHC program indicate that many of the initial participants were large multinational firms. Information from World Bank staffers in Belgrade suggest that smaller firms do not find the benefits of the IHC program to be sufficient to cover the costs associated with participation. Anecdotally, the main constraint that limits firm participation is the requirement that firms provide Serbian customs with a bank guarantee. The cost of financing this guarantee can be unaffordable for small firms. Another

---

[14] We drop three firms that take up the IHC program within 2013, but not before August. These firms have too little data to allow an evaluation. Additionally, 10 firms that adopted IHC during 2011-2013 but did not import continuously after adoption are also excluded. All of the IHC firms, their control firms and placebo firms import continuously throughout the time period that is relevant for the estimation of the impact on each IHC firm.

constraint is that smaller firms' imports are often consolidated within the same shipment as those of other firms. Consolidated shipments are not eligible for clearance under the IHC program.

Table 1. Requirements for Qualification to Participate in the In-House Clearance Program

Each company applying for the process must:
- Have a registered office in the customs territory of the Republic of Serbia,
- Regularly report the goods for the import process, and make an economic justification for adoption of the IHC program;
- Have an authorized person who electronically submits the documents for customs clearance, and keeps records of the goods in the electronic form. This person must have passed the examination for customs broker;
- Establish a control system of engaged carriers in terms of strict rule compliance;
- Designate, in its internal instructions, the exact place within premises where the means of transportation of the declared goods will be located;
- Agree in writing that the company will pay for the expenses incurred for transportation costs of customs representatives leaving the working place and areas in which they regularly performs customs clearance and overtime for customs officers in carrying out the required procedure.
- Enclose with the application a suitable bank guarantee or other form of security.
- Meet other conditions and criteria specified in the regulation.

Source: Information provided by Mr. Nenad Popadic (IFC) on the basis of Serbian Customs Law.

Section 2.3 Hypothesized Impacts

The theory of change to be tested in this study posits a direct effect of the program on customs clearance times and inspection rates. These effects, along with others, could lead participating firms to experience lower (unobserved) trade costs. Lower trade costs could lead participating firms to increase their trade, relative to non-participating firms.

As noted above, the default customs clearance time for shipments clearing under the IHC program will be exactly 0.5 hours. The distribution of clearance times under normal clearance procedures is heterogeneous so firms that adopt the program will see a mechanical compression of the distribution of clearance times at the lower end of the distribution. But participating firms only clear about half of their imports under the program, so it remains unclear how the program will affect their overall distribution of clearance times. Summary statistics shown in Section 4 indicate that 0.5 hours is a reduction, relative to the median clearance time in 2010, for firms that will subsequently adopt the program. The program thus might be expected to reduce median clearance

times. It may or may not affect clearance times above the median, but if some import shipments that would have been cleared much more slowly clear at default levels the program might also reduce clearance times in the upper half of the distribution. We therefore look to see how the program affects clearance times at the 50th, 75th and 90th percentiles of the monthly distribution of customs clearance times. The distribution of clearance times is strongly right skewed, so reductions in the upper half of the distribution act to reduce both the mean and variance of clearance times. We therefore interpret reductions in clearance times at the 75th and 90th percentiles as reductions in firms' uncertainty about clearance times.

There are three complications that may mean individual importing firms do not see a statistically significant reduction in their clearance times as a result of the program. First, even prior to the adoption of the program many shipments clear in less than half an hour. Clearance at the automatic default of one half hour would represent an increase in clearance times for those shipments. Firms that are clearing most if not all shipments in under 0.5 hours before the program could see an increase in their clearance times. Second, even among IHC-adopting firms, many shipments are not cleared under the IHC program but instead go through the normal clearance procedures. If these transactions remain in their place within the distribution of clearance times and are unaffected by the program then one may not see reductions in clearance times in the upper half of the distribution. A third possible source of complication for evaluation are general equilibrium effects. The purpose of the IHC program is to relieve congestion at the customs office by allowing IHC firms to clear remotely. If the adoption of the program by IHC firms reduces clearance times of the control firms by relieving congestion, the estimation procedure may not find an impact on customs clearance time. While this is possible, it seems unlikely. Imports cleared under the IHC program in 2013 account for only 2.6% of Serbia's total value of imports. The channels through which general equilibrium effects might complicate evaluations of the impact on inspections or firm import values are less direct. In any case, at the low levels of IHC take-up observed in our data, it would seem that general equilibrium effects are not likely to be large enough to have a major impact on our findings.

It is also possible that participation in the program would reduce the rate of inspections that IHC-adopting firms face. Shipments cleared under the IHC program are only subject to random inspection, which is rare. For this reason we might expect the program to reduce the inspection rate. Of course, firms continue to clear goods through normal channels after adopting the program.

One reason could be that the transactions are not eligible to be cleared under the program. One might expect higher rates of inspection on such transactions (which would be those deemed most risky by the customs agency). Thus, there are reasons to believe that the inspection rate will not fall all that much as a result of the IHC program.

The links between the IHC program, customs clearance times and inspection rates are relatively straightforward, but a more interesting potential link is effects of the program on firm-level trade volumes. An emerging literature on the time costs of trade suggests that decreases in clearance times or uncertainty about them should reduce (implicit) trade costs, and reduced trade costs should increase trade.[15] Reduced inspections would offer a different, but related, channel. Indeed the IHC program may increase trade through channels other than those we study here, such as reduced transit times from the border to warehouse or other less tangible benefits. Because our sample has relatively short time horizons we are not able to see the long-run effects of reduced trade costs, but the existing literature on customs clearance times has found evidence that increased trade coincides with reduced clearance times rather than lagging them.[16]

One assumption that will be implicit in our identification strategy is that the firm's decision to adopt the IHC program does not depend on anticipation about future changes in the outcome variable. For example, if IHC firms adopted the program because they expected more trade in subsequent periods, we might falsely conclude that the program caused more trade. While anticipation effects are quite plausible for trade volumes, we do not in fact find evidence of a statistically significant impact of the IHC program on trade volumes. Anticipation effects are much less plausible for clearance times or inspection rates. Firms that were expecting lower clearance times and/or fewer inspections in the near future should actually be *less* likely to adopt the program. If anything this would bias our results against finding a negative effect of the IHC program on clearance times or inspection rates.

---

[15] See, especially Hummels and Schaur (2013), Volpe et al. (2015) and Fernandes et al. (2015) on the effect of time reductions on trade. Clark et al. (2013) demonstrate that uncertainty over ship arrival times reduces trade. In most of these studies, units of time are measured in days rather than minutes or hours. In our analysis we will measure time in units as precise as one second. Rather than investigate changes in higher-order moments of clearance time (i.e. uncertainty), we investigate changes at different points of the clearance time distribution. This decision is motivated by the right-skewed clearance time distribution and by the view that it is the relatively longer clearance times that are the primary source of trade costs that operate through uncertainty.

[16] In analyses of risk management programs, both Volpe at al. (2015) and Fernandes et al. (2015) find that time reductions due to reduced inspections increase trade within the same year as the reduced inspections.

3. Estimation Methods

The relatively small number of participating firms in the IHC program, and the particular constraints that limit participation by other firms, imply a situation in which it is difficult to identify a suitable control group of firms using conventional impact evaluation methods. The large scale of imports by participating firms means that the vast majority of firms are not suitable individually as controls. If other large firms differ from participating firms in variables such as the commodity composition of imports, they may not be suitable as plausible controls. More generally, the complexity of customs environments - where thousands of products are traded intermittently by thousands of firms - makes econometric identification extremely difficult. The synthetic control method that we use has some strengths that are useful for this application to the IHC program and, we believe, for other trade facilitation reforms. First, it offers a framework for evaluation when there are a relatively small number of treated units. Second, it is capable of accounting flexibly for multiple unobserved time-varying factors, including trends, seasonality, or other more complex dynamics. Third, the selection of suitable controls for treated units is data-driven, not subject to researcher subjectivity. Finally, the nonparametric structure of the estimator allows for flexible interpretations of lagged responses of the outcome variable to the treatment.

Because the treatment effects that we estimate are modest, relative to the overall volatility of the outcome variables we study, the statistical power of an individual application of the synthetic control method is weak. This problem can be partially overcome in our study because we have multiple treated units. We apply a recently developed method for pooling across estimates from a synthetic control procedure to evaluate joint hypotheses regarding the effect of the IHC program on 21 Serbian importing firms. Pooling allows us to conduct higher-powered tests for program impact, and to test for heterogeneous treatment effects across IHC firms. Finally, the method generates estimates of the common treatment effect that are more robust to outliers than the estimates from other approaches.

3.1 The Synthetic Control Method

In order to be able to select a credible comparison group for the IHC firms, and to address the issue of unobserved time-varying firm heterogeneity, we apply the synthetic control method (SCM) proposed by Abadie and Gardeazabal (2003) and Abadie et al. (2010) to firms adopting the in-house clearance program. SCM is often applied in case studies involving policy interventions

affecting a small number of (large) individual units such as regions or countries. Abadie and Gardeazabal (2003) apply SCM to estimate the economic effects of conflict in the Basque region of Spain while Abadie et al. (2010) apply it to evaluate the impact of a smoking ban in California. In both studies, only one region is treated and the region is unique in important ways that make it difficult to identify a single unaffected region to serve as a control. SCM constructs a 'synthetic' control region as a weighted combination of untreated regions that resemble the characteristics of the treated region before the policy intervention. Then, the evolution of outcomes for the synthetic control region is compared to the evolution of actual outcomes for the treated region after the policy intervention to measure the impact of the intervention.[17]

Below we present the SCM applied to a single IHC firm - as if it there was only one participating firm. Pooled inference is discussed in Section 3.2, which describes how to obtain the overall impact of the IHC program pooling across all treated firms. The methodological descriptions follow closely Dube and Zipperer (2015).

Suppose that we have data for J importing firms with $j=1, 2, \dots, J$ over T periods (months), where $j=1$ is the single treated IHC firm and the other $J-1$ firms are the "donor pool" of potential control firms that are not participating in the IHC program. Let Y_{jt} be an outcome variable of interest and let the sample include a number T_0 of pre-IHC program periods and a number T_1 of post-IHC program periods for a total number of $T = T_0 + T_1$ periods. IHC program adoption is denoted as the indicator variable D_{jt} taking the value of 1 for the treated firm $j=1$ after it adopts the IHC program in period T_0 and 0 otherwise: $D_{jt} = \begin{cases} 1 \ if \ j = 1 \ and \ t > T_0 \\ 0 \qquad otherwise \end{cases}$. The SCM assumes a data generating process whereby observed outcome variable Y_{jt} of firm j in period t is given by:

$$Y_{jt} = \alpha_{jt} D_{jt} + Y_{jt}^N \tag{1}$$

which is the sum of the effect from the treatment $\alpha_{jt} D_{jt}$ and the counterfactual outcome variable (without the IHC program) Y_{jt}^N. Eq. (1) can be written as:

$$Y_{jt} = \alpha_{jt} D_{jt} + \boldsymbol{\theta}_t \boldsymbol{X}_j + \lambda_t \boldsymbol{\mu}_j + \boldsymbol{\delta}_t + \varepsilon_{jt} \tag{2}$$

---

[17] Recent applications of the SCM to trade liberalization episodes by country, to a policy intervention promoting tourism in a region in Argentina, and to the impact of consumer boycotts on trade are given, respectively, by Billmeier and Nanncini (2013), Castillo et al. (2015), and Heilmann (2016).

where $\boldsymbol{\delta}_t$ is a vector of time-specific effects that are common across all firms, $\boldsymbol{X}_j$ is a vector with k observed covariates/characteristics for each firm, $\boldsymbol{\theta}_t$ is a vector of time-varying unknown parameters, and ε_{jt} is an independent and identically distributed (i.i.d.) transitory shock.

The other crucial term included in Eq. (2) is the product of a vector of unknown time-varying factors, $\boldsymbol{\lambda}_t$, and unknown factor loadings, $\boldsymbol{\mu}_j$. These factor loadings can vary across firms, which means that the outcome variables associated with treated and individual control firms are not assumed to follow parallel trends, conditional on observed covariates.[18] If the true factor loadings for the treated firm and for control firms were observable, an unbiased control to serve as counterfactual could be based on the control firms whose factor loadings average to $\boldsymbol{\mu}_1$. Since factor loadings are not observable, the SCM obtains a vector of weights $\boldsymbol{W}$ over control firms in the donor pool such that the weighted combination of those control firms - the synthetic control firm - matches as closely as possible the treated firm in pre-treatment outcomes.

To be more specific, let us define for the IHC treated firm a vector of pre-treatment characteristics, which are covariates and L linear combinations of pre-treatment outcomes, as: $\boldsymbol{Z}_1 = \left(\boldsymbol{X}_1', Y_1^{K_1}, \dots, Y_1^{K_L}\right)$.[19] An analogous matrix with similar variables for the J control firms is denoted by $\boldsymbol{Z}_0$. The SCM chooses a vector of weights $\boldsymbol{W}$ to attach to control firms in the donor pool so as to minimize the distance between pre-treatment characteristics of the IHC firm $\boldsymbol{Z}_1'$ and of control firms $\boldsymbol{Z}_0'$. That distance equals the mean square prediction error over k pre-treatment characteristics:

$$\sum_{m=1}^{k} v_m (\boldsymbol{Z}_{1m} - \boldsymbol{Z}_{0m}\boldsymbol{W})^2 \tag{3}$$

where v_m reflects the relative importance assigned to the mth predictor in vector $\boldsymbol{Z}$. Given optimal weights w_j* for each of the $j=2, \dots, J$ control firms, the synthetic control firm in period t (the estimator for the counterfactual outcome Y_{jt}^N) is the weighted average of the outcome variable for the control firms: $\sum_{j=2}^{J} w_j^* Y_{jt}$. The unbiased SCM estimator of the impact of the IHC program for IHC firm 1 in post-treatment period $t = T_0 + 1, \dots, T$ is given by comparing the evolution of the

---

[18] Importantly, this is very different from the requirement of the traditional alternative method, the DiD regression. The DiD estimator assumes that treated and control firms' outcomes would have followed similar time trends in the absence of treatment (conditional on differences in outcome levels). From Eq. (2) a DiD estimating equation can be obtained if unobserved firm heterogeneity is *not* allowed to vary over time, i.e., if $\boldsymbol{\lambda}_t$ is constant over time. Two sets of differences would cause $\boldsymbol{\mu}_j$ and $\boldsymbol{\lambda}_t$ to drop out of the regression with $\boldsymbol{\lambda}_t$ constant.
[19] Dube and Zipperer (2015) argue that the unbiasedness of the SCM relies on those linear combinations of pre-treatment outcomes being included in the vector Z of predictors of the outcome.

outcome variable for the synthetic control firm to the evolution of the same variable for the IHC firm as in:

$$\widehat{\alpha_{1t}} = Y_{1t} - \sum_{j=2}^{J} w_j^* Y_{jt} \tag{4}$$

Abadie et al. (2010) show that for sufficiently large T_0 the weights w_j^* are such that the synthetic control firm has average factor loadings that match the loadings of the treated firm, that is, they satisfy $\sum_{j=2}^{J} w_j^* \mu_j = \mu_1$.[20] This means that the underlying dynamics of the outcome variable that are particular to the IHC firm are also present in the synthetic control firm. The SCM estimator's ability to remove the influence of the $\lambda_t \mu_j$ term in Eq. (2) gives it an advantage over a DiD estimator in this and many other settings involving customs reform. In particular it means that the procedure can account for program participation that is endogenous to the pre-treatment time path of the outcome variable.[21]

A key identifying assumption in the SCM is the independence of treatment status and potential outcomes, conditional on pre-treatment observable covariates and lagged outcomes. While this assumption is similar to that required of the DiD estimator when combined with propensity score matching proposed by Heckman et al. (1997), the difference is that the SCM's matching on pre-treatment values of outcomes and relevant covariates helps to control for observed and unobserved determinants of those outcomes and for the heterogeneity of the effect of those covariates in order to estimate a causal effect of the program.

From Eq. (3) it is clear that the weight given to each pre-treatment characteristic used as predictor, v_m, influences the choice of weights w_j^* on control firms that determine the synthetic control firm. Optimal weights w_j^* are determined by solving:

$$W^*(V) = argmin_W \sum_{m=1}^{k} v_m (Z_{1m} - Z_{0m} W)^2 \tag{5}$$

We follow Abadie et al. (2010) in a data-driven selection of the optimal V^* such that the synthetic control firm best fits pre-treatment outcomes of the treated firm.[22] Specifically, the optimal V^* is

---

[20] The key identifying assumption that T_0 is sufficiently large is required so that any correlation between the weights w_j and the error terms goes to zero. In our case $T_0 = 19$, which is roughly similar in scale to those used in other studies (such as Abadie et al. 2010), although our units are months rather than years.

[21] At one extreme, λ_t could be a vector of time fixed effects, so its product with heterogeneous firm weights μ_j would be equivalent to firm-specific time-varying fixed effects. Simpler dynamics where λ_t represents a cyclical and a trend component would have the matrix μ_j representing firm-varying exposure to such factors. For these and many other representations of dynamics in the outcome variable, the SCM removes the role of $\lambda_t \mu_j$.

[22] Abadie et al. (2010) and Abadie (2014) provide a detailed discussion on the methods to choose V.

the solution to the joint nested optimization problem given by Eq. (5) and the equation below that minimizes the fit in the pre-treatment outcome variable:

$$\boldsymbol{V}^* = argmin_V \sum_{t=1}^{T_0}\left(Y_{1t} - Y_{0t}\boldsymbol{W}(\boldsymbol{V})\right)^2 \tag{6}$$

An iterative solution to (5) and (6) produces jointly optimal weights $\boldsymbol{W}^*(\boldsymbol{V}^*)$. Note that in this optimization two constraints are imposed on the weights w_j^*: that they are all non-negative and that they sum to 1. These two restrictions constrain the estimator from extrapolating outside the donor pool.

A related question that arises is that of which conditioning variables to include in vector $\boldsymbol{Z}$. As Dube and Zipperer (2015) mention different sets of predictor variables may result in different synthetic control firms but there is little guidance in the econometric literature to assess predictor choice. Our approach will be to present benchmark results using our preferred set of predictors but to also show robustness results where the set of predictors is allowed to vary. Our preferred benchmark estimates are those chosen by the Dube and Zipperer criterion.

In sum, SCM estimates the missing counterfactual for IHC firm 1 as a weighted average of pre-treatment outcome variables and conditioning variables associated with the donor firms identified by the procedure. The weights are chosen such that the resulting synthetic control firm has covariates and pre-treatment outcomes that are on average very similar to those of IHC firm 1. Then, if post-treatment IHC firm 1 and its synthetic control firm exhibit a discrepancy in outcomes, the discrepancy is interpreted as having been caused by the IHC program.

An attractive feature of studies using the SCM is the transparency of the estimator. In studies in which the treated unit is a region, a description of the numeric weights w_j^* attached to the donor units can be informative for the reader.[23] In our case confidentiality agreements with the Serbian customs authority mean the treated firms and the donor firms must remain anonymous. Studies of single units often also report the components of the **V** matrix, which helpfully identify the characteristics that are most useful in fitting pre-treatment outcomes. We will forgo that in our study because the **V** matrices differ in our many applications of the method. The application of the SCM to the evaluation of the IHC program still exploits the other attractive features of the method, such as the safeguard against extrapolation, flexibility in terms of potential restrictions

---

[23] For example in Abadie et al. (2010) the goal is to construct a synthetic California based on states that did not adopt a tobacco control program during the sample period. From the reported results we learn that the synthetic California produced by the estimating procedure is a weighted average of the outcome variables for Utah, Nevada, Colorado, Montana and Connecticut, with weights 0.334, 0.234, 0.199, 0.164, and 0.069, respectively.

that can be imposed on the donor pool, and the fact that unobservable confounding firm heterogeneity is allowed to vary over time.

3.2   Placebo Tests, Statistical Inference and Pooling across Treated Firms

Once the difference between the IHC firm's outcome and that of its synthetic control firm is calculated for different periods (as in Eq. (4)), it is necessary to determine whether the estimated effects are statistically significant. To that end, we implement the inferential techniques of placebo or falsification tests based on permutation techniques proposed by Abadie et al. (2010) and pooling techniques proposed by Dube and Zipperer (2015).

Statistical inference for a single treated firm can be conducted by generating a distribution of outcomes that arise from applying the SCM to placebo firms. Most applications of the SCM have a relatively small set of potential donors so their placebo firms are drawn from the donor pool used to construct a synthetic control for the treated unit. Our potential donor pool is large as it is the universe of importing firms in Serbia, which allows us to exclude the placebos from the donor pool. From within the same decile of firm total import value in 2010, we randomly select P = 25 untreated donor firms to be the placebo firms and remove them from the donor pool that is used for the construction of the synthetic control firm. The placebo firms use the same donor pool as the treated firm but they are ineligible to be donors themselves in the construction of the synthetic control firm.

For each IHC firm, we iteratively assign 'treatment' to each of its placebo firms, treating them as if they began using IHC in the same month as the treated firm. That is, the procedure described in Section 3.1 is carried out 25 times, taking in turn each of the placebo firms to be the 'treated' firm, j = 1. For the placebo firm P a synthetic placebo firm is identified following the same method as for treated firms and 'treatment' effects $\hat{\alpha}_t^P$ (defined as in Eq. (4)) are estimated for each period t. Because the placebo firms were randomly chosen and not exposed to the IHC program, the distribution of the $\hat{\alpha}_t^P$'s for any given period t defines a simulated distribution for the treatment effect $\hat{\alpha}_{jt}$ under the assumption that the impact of the IHC program was exactly zero. For a single treated firm informal inference can be conducted by plotting the time path of $\hat{\alpha}_{jt}$

against the corresponding $\hat{\alpha}_t^P$'s. If $\hat{\alpha}_{jt}$ lies well inside the distribution of the simulated $\hat{\alpha}_t^P$'s, there is very little evidence to conclude that the IHC program affected firm j's outcome Y.[24]

For the purposes of formalizing testable hypotheses, and for measuring the impact of the IHC program, the point estimates of the treatment effects on treated firm f, the $\hat{\alpha}_{ft}$'s, vary monthly and thus represent a moving target. (Notice that we will use the f subscript to designate the 21 IHC-treated firms in exercises that involve pooling over treated firms.)[25] The flexibility of point estimates is helpful for descriptive purposes but not helpful for quantifying impact. We propose an additional statistic following Dube and Zipperer (2015) that identifies a particular window of time over which the impact of the IHC program will be assessed. Formally, we construct a statistic that averages the point estimates of the treatment effects for IHC firm f (the $\hat{\alpha}_{ft}'$s) over $T - T_0$ post-treatment periods (months) as follows:

$$\hat{\beta}_f = \frac{1}{(T-T_0)}\Sigma_{t=T_0+1}^T(Y_{ft} - \Sigma_{k=2}^{K^f} w_k Y_{kt}) \tag{7}$$

where K^f is the donor pool for treated firm f. We use $\hat{\beta}_f$ as the main point estimate of the treatment effect of the IHC program for firm f. In our main specification and in several of the robustness checks we will use a 6-month window, $T - T_0 = 6$. One robustness check will employ a 9-month window, $T - T_0 = 9$, at the cost of a smaller sample size. These window lengths are limited by our data's sample period and the fact that firms take-up the IHC program on different dates.

For the placebo test, point estimates of the impact of the IHC program are calculated as in Eq. (7) for each placebo firm, and are designated as $\hat{\beta}^P$. One way to ask if the IHC program had a statistically significant effect on firm f's outcome Y over $T - T_0$ post-treatment periods is to ask whether the $\hat{\beta}_f$ statistic lies in the extreme tails of the distribution of the corresponding placebo $\hat{\beta}^P$ statistics. To answer this question we rank each treated firm's point estimate of the IHC treatment effect in relation to the point estimates of its 25 placebo firms.[26] We define the treated firm's percentile rank statistic as the rank of its treatment effect estimate relative to the 25 placebo firms divided by 26, which is the total number of firms (25 placebos plus one treated). The idea

---

[24] In Appendix C we show an example for a single treated firm. In that case the values of $\hat{\alpha}_{jt}$ for log median clearance times are clearly negative, and towards the bottom of the distribution of the $\hat{\alpha}_t^P$'s, but not clearly outside that distribution.

[25] The j subscript used before denoted the set of firms used in the analysis of a single treated firm ($j=1$).

[26] The underlying distribution of β's is not fully revealed by a simulation of 25 values under the null hypothesis. But the rank order of $\hat{\beta}_f$ relative to the set of $\hat{\beta}^P$'s is approximately uniform under the null hypothesis, and this allows formal hypothesis testing of small samples under the null hypothesis. This is the key rationale for using rank order tests.

behind this placebo test is to measure whether the effect estimated by the SCM for an IHC-adopting firm is large, relative to the effect that would be estimated for a firm taken at random. The placebo firms summarize outcomes under the null hypothesis of no treatment effect.

More formally, to summarize the position of the treatment effect statistic $\hat{\beta}_f$ for treated firm f, relative to the simulated distribution under the null, we calculate the percentile rank statistic $p_f = \hat{G}_f(\hat{\beta}_f)$ where $\hat{G}_f$ is the empirical cumulative distribution function (CDF) of the $\hat{\beta}_f$ statistic. As Dube and Zipperer (2015) note, the percentile rank statistic is (approximately) uniformly distributed on the unit interval under the null hypothesis of no effect of the IHC program. That is, under the null hypothesis of $\hat{\beta}_f = 0$, the values of p_f should be distributed uniformly as if the values of $\hat{\beta}_f$ were generated by random noise. For a single treated firm our use of 25 placebo firms implies that we can only claim statistical significance - at the 8 percent confidence level, for a two-sided test - if the percentile rank statistic lies outside of the distribution of the simulated values of $\hat{\beta}_f$ for the placebo firms.[27]

Estimates for a single treated firm lack statistical power in this setting, but statistical inference can be strengthened by pooling across IHC firms. In order to measure the mean treatment effect we use the simple average of the treatment effects $\hat{\beta}_f$ (calculated as in Eq. (7)) across F treated firms: $\bar{\beta} = \frac{\sum_{f=1}^{F} \hat{\beta}_f}{F}$. But for purposes of hypothesis testing it is more useful to rely on the mean of the percentile rank statistic p_f across the F treated firms: $\bar{p} = \frac{\sum_{f=1}^{F} p_f}{F}$. As indicated by Dube and Zipperer (2015), the mean percentile rank statistic $\bar{p}$ follows a Irwin-Hall distribution, as it is the mean of F uniformly distributed random variables under the sharp null hypothesis of no effect of the IHC program on any treated firm. Use of statistic with a known distribution facilitates hypothesis testing even when the number of treated units is small.[28]

Following Dube and Zipperer (2015), we can use a simulated Irwin-Hall distribution to calculate confidence intervals for given values of F and $\bar{p}$ under the null hypothesis of no effect of

---

[27] Dube and Zipperer (2015) use 20 placebo firms, which gives them a more natural link to hypothesis testing (the implied p-value is 0.10 rather than 0.08). Because we have 21 treated firms we found that readers had difficulty distinguishing calculations involving an individual treated firm's rank within its distribution of 20 placebo firms from calculations that are done across the 21 treated firms. Since we are not interested in tests at the level of an individual firm, we use 25 placebo firms rather than 20, which helps to avoid this specific source of confusion. A larger set of placebos also raises the precision of our p_f estimates.

[28] By contrast the distribution of the $\hat{\beta}_f$ statistic is not known, and would be difficult to simulate credibly for small numbers of treated units.

the IHC program on any treated firm.[29] Based on the percentiles of the distribution of the mean of F=21 uniformly distributed random variables provided in Table A.1 of Dube and Zipperer (2015), a two-sided test of the null hypothesis of no impact of the program can be rejected at a 5 percent significance level if the mean percentile rank statistic, $\bar{p}$, falls below 0.377 or above 0.623.

Dube and Zipperer (2015) suggest several advantages in conducting pooled inference in this manner. Of these, the most relevant are that 1) tests involving the percentile rank statistic are a natural generalization of the placebo tests conducted in SCM studies with a single treated unit, 2) the mean of the percentile ranks has a known distribution under the sharp null hypothesis (allowing for exact inference), and 3) the use of the percentile rank statistic $\bar{p}$ diminishes the impact of outliers, relative to tests that involve the point estimates $\hat{\beta}_f$. A limitation of the test is that it is an evaluation of the sharp null hypothesis that the effect of the IHC program is zero for *all* treated firms, rather than a test that the mean effect of the IHC program across treated firms is zero. This limitation can be addressed by testing for heterogeneous treatment effects, as described in Section 3.4.

3.3 Calculating Hodges-Lehmann Statistics Using Inverted Rank Statistics

Comparisons of estimated mean treatment effects ($\bar{\beta}$) and median treatment effects of the IHC program on a given outcome in Section 5 will demonstrate that the estimated mean treatment effect is affected by outliers. While p-values associated with the mean of the percentile rank statistic can provide information about the statistical significance of the estimate, they do not offer much guidance about the range of plausible treatment effects. Dube and Zipperer (2015) propose the use of inverted percentile rank statistics as formulated by Hodges Jr and Lehmann (1963) for characterizing the 95% confidence intervals for the mean treatment effect. The confidence interval defines the set of critical values β for which the estimated mean treatment effect, $\bar{\beta}$, cannot be rejected when those values of β are chosen as the null hypothesis.

Formally the 95% Hodges-Lehmann confidence interval is calculated by inverting a mean-adjusted percentile rank statistic at the desired critical values using the empirical distribution of treatment effects for the placebo firms. Following Dube and Zipperer we define the mean-adjusted percentile rank statistic as:

---

[29] The Irwin-Hall distribution converges to a standard normal for large values of F. But since we are testing for relatively small values of F, it is quite useful to have a formulation for the exact distribution.

$$\overline{p}(\tau) = \frac{\sum_f \hat{P}_f (\beta_f - \tau)}{F} \qquad (9)$$

where $\hat{P}_f(.)$ represents the firm-specific percentile rank after the firm-specific treatment effect $\hat{\beta}_f$ estimate is adjusted by the parameter τ. Let $G_f(\overline{p}, F)$ be the CDF of the mean percentile rank statistic of F uniformly distributed variables, and $\hat{G}_f(\overline{p}(\tau), F)$ the empirical counterpart for a given adjustment τ. The boundaries of the 95% Hodges-Lehmann confidence interval are calculated by identifying the values of τ such that $\hat{G}_f(\overline{p}(\tau), F) = 0.025$ and $\hat{G}_f(\overline{p}(\tau), F) = 0.975$. One can also generate a meaningful point estimate of the common treatment effect - the Hodges-Lehmann point estimate - by calculating τ such that $\hat{G}_f(\overline{p}(\tau), F) = 0.5$. This is the value of the adjustment τ that would put the mean percentile rank statistic precisely at 0.5, given the empirical distribution of estimated treatment effects for both treated and placebo firms.[30]

We adopt the Hodges-Lehmann point estimate as our preferred measure of central tendency for the treatment effect, although we also report estimates of the mean and the median treatment effect. While the latter estimates are more familiar, the Hodges-Lehmann estimate is less sensitive to the idiosyncrasies that arise in finite samples because it depends on ranked positions. The Hodges-Lehmann confidence interval is also quite useful in small samples because it offers a precise sense of the confidence interval surrounding the estimate of the common treatment effect.

3.4 Testing for Heterogeneous Impacts

The test of the mean percentile rank statistic that Dube and Zipperer (2015) propose for the pooled estimates has important strengths from the point of view of statistical inference, but it evaluates a particular null hypothesis, that there is no impact of the IHC program on the outcome variable for *any* treated firm in the sample. DiD regressions, by contrast, would test the hypothesis that the estimated mean treatment effect is different than zero.[31] Insofar as our interest lies in the joint effect of the IHC program across adopting firms, the mean percentile rank statistic test is not fully adequate.

---

[30] Dube and Zipperer (2015) provide more detail on the calculation of the Hodges Lehmann confidence intervals and point estimates and a graphical representation of the calculation in Figure 1 of their paper.

[31] As mentioned in Section 3.2, testing a hypothesis about the mean treatment effect is difficult in our setting because the distribution of the mean treatment effect is difficult to calculate due to the a) small sample of treated units and b) the potential for heterogeneous treatment effects across firms.

To address this issue Dube and Zipperer (2015) propose subsequent tests that evaluate the heterogeneity of the estimated treatment effects across firms. As with the mean percentile rank statistic test, the proposed Kolmogorov-Smirnov and Anderson-Darling tests rely on the uniform distribution of the percentile rank statistic under the null hypothesis of no impact of the program on any treated firm. The heterogeneity tests, however, do not focus on properties of the mean percentile rank statistic *per se*, but rather on the entire distribution of the p_f statistic. Thus rather than evaluating the hypothesis that the mean percentile rank statistic differs from zero, these tests ask whether the entire distribution of percentile rank statistics differs from the uniform distribution implied by the null hypothesis. Under the null hypothesis, the cumulative distribution of an ordered percentile rank statistic should approximate an upward sloping line (or for small samples, a step function approximated by an upward sloping line). Dube and Zipperer propose to test these hypotheses by ordering the percentile rank statistics and applying the Kolmogorov-Smirnov and Anderson Darling tests against the presumed null hypothesis of a uniform distribution.[32]

The heterogeneity tests can be used to make two types of inferences. First, if the mean percentile rank statistic test indicates a significant treatment effect, heterogeneity tests can be used to make inferences about the commonality of the estimated treatment effect across treated firms. Dube and Zipperer propose a test in which all treated firms' estimated treatment effects $\hat{\beta}_f$ are adjusted by subtracting the mean treatment effect $\bar{\beta}$, and the resulting adjusted $\hat{\beta}_f$ are ranked relative to the distribution of effects for the placebos. This implies an adjusted percentile rank statistic for each treated firm, centered around the mean effect. If the distribution of the adjusted percentile rank statistics is not significantly different from a uniform distribution, then we cannot reject the hypothesis of a treatment effect that is common across firms.

Second, if the mean percentile rank statistic test on unadjusted data does not reject the null hypothesis of no treatment effect, the absence of an observed effect could be due to the presence of offsetting treatment effects of different signs occurring on either side of the mean effect. In the case of the IHC program, such an outcome is most likely with respect to the variable measuring the change in firms' median customs clearance times. Because the IHC program has a default clearance time of 0.5 hours for shipments that are not subject to inspection, the program can

---

[32] The Kolmogorov-Smirnov is a non-parametric test for the equality across two probability distributions while the Anderson-Darling is a test of whether a given sample of data is drawn from a given probability distribution. Both tests are explained at greater length in Dube and Zipperer (2015). The Anderson-Darling test is more sensitive to extreme values than the Kolmogorov-Smirnov test.

potentially increase median clearance times for some firms while reducing it for others.[33] To examine the presence of offsetting effects on either side of the mean treatment effect we compare the distribution of the unadjusted (original) percentile rank statistic to a uniform distribution. In Section 5, we apply both mean-adjusted and unadjusted heterogeneity tests to all of our outcome variables, using the results from our benchmark specification of the SCM.

4. Data, Summary Statistics and Implementation

4.1. Data

The raw data we employ track Serbian imports for consumption at the transaction level for the period 2010-2013. The data were provided confidentially by the Serbian national customs authority. The Serbian customs agency processes goods at the level of individual import declarations (each declaration may contain one or more transactions). The primary fields of interest used in this study are the registration time and the clearance time for each declaration (which allow us to calculate the clearance time in hours per declaration), a numeric identifier for the importing firm, a Harmonized System (HS) 10-digit commodity code for each transaction, the country of origin, a field indicating what type of inspection the declaration was subject to, and a field indicating whether the declaration was cleared under a special customs clearance procedure. This latter code which we designate as 'special clearance' includes transactions of goods that were cleared under the IHC program, but also includes other special clearance codes.[34]

Our outcome variables of interest, calculated from the raw data aggregated to the firm-month level, are a) the (log) median time necessary for the firm to clear customs for its import shipments in a month, b) the (log of the) 75th percentile of customs clearance times in a month, c) the (log of the) 90th percentile of customs clearance times in a month, d) the rate of inspections of the firm's import declarations in a month, and e) the (log) value of imports of the firm in a month. We use locations in the distribution of clearance time (percentiles), rather than moments of the

---

[33] Firms that routinely clear their goods in less than 0.5 hours prior to adopting the IHC program could see increases in their clearance times after adoption of the program, while firms with customs clearance times lower than 0.5 hours would see reductions. These offsetting effects across firms might lead the mean percentile rank statistic test to accept the null hypothesis of no effect of the IHC program on any firm. The tests for heterogeneous treatment could at the same time find evidence of program impact.

[34] Other than the code that identifies transactions cleared under the IHC program, we are not privy to the meaning of these codes. Our use of a generic special code dummy presupposes that those codes offer information about firm types that is useful for matching treated and donor firms in the pre-treatment period.

distribution, and we apply the log function in order to reduce the month-to-month volatility.[35] Volatility greatly reduces computational speed as well as the performance of the SCM estimator, so we follow previous studies in pre-treating the data in ways that reduce the volatility of the outcome variables. Inherent variability in clearance times and in firms' monthly import activity generates a high degree of volatility in the outcome variables even in logs. We further smooth the fluctuations in the outcome variables by calculating the 3-month moving average (i.e., the average of the current and previous two months) for each of the outcome variables a)-e) for both treated and control firms. These moving averages are the actual outcome variables we use in the impact evaluation of the IHC program.[36] These variables nonetheless remain more volatile than is common in most SCM studies.

The SCM uses a weighted combination of non-IHC firms to construct a synthetic control firm that resembles in the most relevant economic characteristics the IHC firm before the outset of the IHC program. We rely on the administrative trade data to define the vector of conditioning variables that are used to construct the synthetic control firms. For each firm and month we calculate the share of imports originating in the EU, the share of imports cleared under a special customs clearance code (other than IHC), and the share of imports in each of 10 broad commodity groups or sectors. Appendix Table 1 provides the detailed definitions of the conditioning variables. In particular, note that the use of sector data allows us to account for seasonal patterns in imports across sectors and for supply or demand shocks that would apply to firms trading in the sector. The vector of conditioning variables also includes the relevant outcome variables in specific pre-treatment periods that are discussed below.

The purpose of the conditioning variables is to ensure that the synthetic control firm not only matches the treated firm in terms of the dynamics of the outcome variable, but also matches important characteristics of the treated firm in the pre-treatment period. The share of imports originating in the EU is useful for conditioning the construction of the synthetic control firm if shocks to the outcome variable are related to the origin of the goods that firms import. The sectoral import shares ensure that the synthetic control firm imports broadly the same commodity mix as the treated firm. A special customs clearance variable matches the synthetic control firm to the

---

[35] The exception is the inspection rate, which we do not log. For IHC-adopting firms, inspection rates are quite low throughout the sample and logging values quite close to zero seems inappropriate for inference.
[36] Transformation of the outcome variable via a moving average has been done in previous applications of the SCM. Cunningham and Shah (2014), for example, smooth their state-level crime and public health outcomes using a moving average before they apply the SCM. Use of the moving average further reduces the role of noise arising from fluctuations in the outcome variable.

treated firm on the basis of their exposure to special import regimes. Ideally our vector of conditioning variables would include a larger number of source region and product categories, but we find that the inclusion of larger numbers of conditioning variables generates computational difficulties. Our selection of conditioning variables allows the procedure to construct synthetic control firms that match the IHC firms along important dimensions, while still retaining an ability to complete the 546 procedures (21 treated firms and 21*25=525 placebos) used to evaluate the impact of the IHC program on each outcome variable within a reasonable time-frame. Details on implementation are provided in Section 4.3.

An important subset of the conditioning variables are the values of the pre-treatment outcome variables that are used to anchor the synthetic control firm's pre-treatment outcomes at approximately the same level as those of the treated firm. In order to check the robustness of our results, we will use different specifications in terms of the lagged outcome variables that are used as conditioning variables. In our benchmark specification we follow Abadie et al. (2010) in using three pre-treatment values of the outcome variables: we choose the values of the (moving averages of the) outcome variable for each firm in the first, 10th and 18th pre-treatment months as the conditioning variables. In one robustness specification we use all pre-treatment values of the outcome variable, thus putting an emphasis on constructing a synthetic control firm that very nearly replicates the pre-treatment dynamics of the treated firm. In another robustness specification we use only the average across all pre-treatment values of the outcome variable, aiding the fitting procedure in replicating the approximate level of the pre-treatment outcome, but giving no direct assistance in hitting the pre-treatment dynamics. We discuss these robustness checks in Section 6 and in greater detail in Appendix C.

Since treated firms began using the IHC program at different dates (from July 2011 onwards) and our sample period covers all months in the 2010-2013 period for each firm, we will balance our sample so that each treated and potential donor firm imports monthly in each of 19 pre-treatment months and in all post-treatment months until the end of 2013.[37] The application of

---

[37] The relatively high degree of month-to-month volatility in the data means that we need as many pre-treatment periods as possible. Moreover, the use of outcomes measured as moving averages consumes even more time periods. The first firm to adopt the IHC program in Serbia does so in July 2011, the 19th month of the sample. For this firm alone we use a 2-month moving average and a pre-treatment period of 17 months. Our sample ends in December 2013, and the number of post-treatment months thus depends on the initial month when each firm uses the program. When we pool across firms we must choose a common time window when to evaluate the program and we will choose 6 post-treatment months in our main specification and 9 post-treatment months in our robustness analysis.

the pooled version of the SCM is more transparent if all IHC firms are given a pre-treatment period of similar length.

4.2. Characteristics of Imports under In-House Clearance and by Qualified Firms

In this section we report some preliminary statistics on Serbia's IHC program with the purpose of understanding how the program is used across products and source countries. We also wish to better understand how firms that will later use the program differ from those that will not use it. To do this we identify all firms that use the program in 2013 and investigate the characteristics of their imports.

Table 2 provides two decompositions of Serbian imports in 2013. Panel A illustrates use of the IHC program by sector. Column 2 shows the share of each sector in total Serbian imports. Minerals is the largest sector, followed by chemicals, plastics and rubbers. Column 3 shows the share of each sector's imports that is imported by firms that adopt the IHC program (whether the imports are cleared under IHC or not). IHC firms are especially large importers of wood and paper products, but import virtually nothing in the transportation sector and are barely active in the minerals sector. Column 4 shows the share of import value (at the sector level) that is imported under the IHC program. These figures differ from those in Column 3 because IHC firms do not clear all their imports under the IHC program. Column 5 shows the share of IHC firms' imports in each sector that were cleared under the IHC program. Firms are most likely to clear animal and vegetable products and foodstuffs, followed by wood and paper products, metals, and raw hides, skins, leathers, textiles and footwear.[38] In all of these sectors, IHC firms cleared more than half their imports under the program in 2013. The final row of Panel A shows that IHC firms accounted for approximately 5 percent of total import value in 2013, and that imports cleared under the IHC program accounted for 2.6 percent of total import value. The key lesson is that IHC firms differ from total imports in their sector profile, that they clear some commodities more often under the IHC program, and that IHC transactions account for a rather small share of total import value in 2013.

Panel B of Table 3 reports imports by source focusing on the differentiation across EU and other sources that will be used in our analysis. The EU was the source of just over half of total

---

[38] A more detailed review of usage at the HS2 level revealed that IHC firms typically used the program to clear their imports of bulk commodities such as wood pulp, wool and other animal hairs, base metals, fertilizers, etc. Tobacco products and sugars and confectionary products are also frequently imported under the IHC program.

import value in 2013. IHC firms accounted for 7.6 percent of EU imports, but only 2.5 percent of imports from other sources. Imports cleared under the IHC program accounted for 3.9 percent of import value sourced in the EU, and 1.3 percent from other sources. IHC firms cleared approximately half of their import value under the program for both source regions.

Table 2. Shares of Total 2013 Import Value and Import Shares

Panel A. By Sector

	Share of total import value	IHC firms' share of sector import value	Share of total import value cleared under IHC	IHC share of IHC firm imports
Animal and vegetable products and foodstuffs	0.101	0.039	0.032	0.812
Chemicals, plastics and rubbers	0.210	0.089	0.033	0.367
Raw hides, skins, leathers, furs, textiles and footwear	0.058	0.054	0.032	0.601
Machines and electrical	0.164	0.034	0.012	0.340
Metals	0.093	0.091	0.057	0.626
Minerals	0.208	0.002	0.000	0.182
Misc. and services	0.039	0.027	0.010	0.384
Stone and glass	0.017	0.053	0.018	0.334
Transportation	0.063	0.000	0.000	0.142
Wood and paper products	0.047	0.194	0.136	0.702
Total	1	0.051	0.026	0.512

Panel B. By Broad Source

	Share of total import value	IHC firms' share of source region import value	Share of total import value cleared under IHC	IHC share of IHC firm imports
European Union	0.518	0.076	0.039	0.514
Other	0.482	0.025	0.013	0.506
Total	1	0.051	0.026	0.512

Source: Authors' calculations based on Serbian customs data.

We now turn to an analysis of 2010 data in order to show that even before the IHC program was undertaken, IHC firms and non-IHC firms differed in the distribution of customs clearance times, in their inspection rates, and in the scale of their imports. Table 3 shows summary statistics on customs clearance times in 2010 for firms that do and do not adopt the IHC program by the end of 2013. The striking pattern that emerges is that the distribution of clearance times for firms that adopt the IHC program by 2013 is well below that of firms that do not adopt, at all locations in the distribution. In particular the firms that will later adopt the program exhibit substantially faster clearance times at the 90th percentile - 3.6 hours versus almost 3 days (70.4 hours) - which indicates

that they have far fewer outlier import declarations that take an abnormally long time to clear customs than do other firms.

Table 3. Customs Clearance Times in 2010 for IHC and Non-IHC Firms

	Distribution of declaration-level customs clearance times in 2010 in hours							
	Mean	25th percentile	Median	75th percentile	90th percentile	Standard deviation	Coefficient of variation	Number of import declarations
IHC firms by the end of 2013	6.66	0.40	0.84	1.72	3.64	45.46	6.83	23,516
Non-IHC firms by the end of 2013	36.16	0.57	1.41	5.03	70.42	440.57	12.18	622,264
All firms	35.09	0.56	1.37	4.66	69.35	432.59	12.33	645,780

Source: Authors' calculations based on Serbian customs data.

Table 4 shows the distribution of monthly inspection rates in 2010 for firms that did and did not adopt the IHC program by the end of 2013. The mean inspection rate for IHC firms is lower than for non-IHC firms, and IHC inspection rates are much lower in the upper half of the distribution. For inspections it is not the case that the distribution for IHC firms is strictly below that of non-IHC firms. Many non-IHC firms are substantially smaller than IHC firms, and so many would import few shipments in a given month. Thus it is possible that even with relatively higher frequency of inspections for non-IHC firms there could be months in which those firms' relatively few shipments were not stopped for inspection. This explains the relatively large share of firm-month combinations with an inspection rate of zero among non-IHC firms.

Table 4. Monthly Inspection Rates in 2010 for IHC and Non-IHC firms

	Distribution of firm-month inspection rates in 2010						
	Mean	25th percentile	Median	75th percentile	90th percentile	Standard deviation	Number of firm-month combinations
IHC firm by end of 2013	0.08	0.04	0.07	0.11	0.17	0.08	381
Not IHC firm by end of 2013	0.20	0	0	0.27	1	0.34	110220
All firms	0.20	0	0	0.26	1	0.34	110601

Source: Authors' calculations based on Serbian customs data.

Table 5 shows summary statistics on the annual import values in 2010 for firms that do and do not adopt the IHC program by the end of 2013. IHC firms are, on average, much larger importers: the mean annual import value for an IHC firm is 27 times higher than for non-IHC

firms. But Serbia's largest importing firm is not an IHC firm and the dispersion in annual import values – as measured by range of the distribution – is actually much wider for non-IHC firms. This is precisely what we want for the purposes of constructing synthetic control firms, namely we would like some very large importing firms that are not participants and some firms that trade the same goods as the treated firm. The SCM uses data from non-IHC firms to construct a synthetic control firm that looks like an IHC firm, but that is possible only if the non-IHC firms demonstrate enough variation in their characteristics that they are capable of reproducing the IHC firms in terms of import composition and dynamics.

Table 5. Annual Import Values in 2010 for IHC and Non-IHC Firms

	Distribution of firm import values in 2010 in Serbian Dinars					
	Mean	Median	Standard deviation	Minimum	Maximum	Number of firms
IHC firms by the end of 2013	1,354 million	654 million	1,656 million	53 million	6,013 million	32
Non-IHC firms by the end of 2013	49 million	2.7 million	881 million	50	93,633 million	22,485
All firms	51 million	2.7 million	884 million	50	93,633 million	22,517
Ratio IHC/Non-IHC	27	244	1.88	1.1 million	0.06	

Source: Authors' calculations based on Serbian customs data.

4.3 Implementation Details

In Section 3 we described the SCM in abstract terms. In practice our application of the SCM faced a number of challenges. We describe here briefly the details of implementation.[39] The primary challenge for estimation was that even the smoothed outcome variables exhibit a high degree of volatility. A different issue (with related implications) is that we use a relatively large number of conditioning variables, especially with respect to the composition of imports.[40] These two features of our analysis mean that we require substantially larger donor pools than is typical in SCM studies, because smaller donor pools were unable to successfully construct the synthetic

---

[39] We perform our estimation using the *synth* routine for STATA developed by Abadie et al. (2010) and use the full nesting procedure to determine the *V* matrix. The procedure identifies through a regression the relative predictive power of the k covariates in the vector Z for the outcome variable and assigns larger weights to covariates with larger predictive power. The procedure takes those weights as initial values for the importance of each covariate and then iterates until optimal weights V* and W* are obtained. The method iterates between solving for the V* matrix and the W* matrix until it reaches an optimal solution. In our case we limit the number of iterations between V and W matrices to 50 in order to achieve a solution in a reasonable period of time.

[40] We could simplify the analysis considerably by aggregating commodities into even larger bundles: primary products and manufactured goods, for example. This would give the estimator fewer targets to hit and so imply a better statistical fit. But more aggregation gives less confidence that the synthetic control firm constructed by the SCM procedure offers a meaningful counterfactual. This is a trade-off but we err on the side of less aggregation in order to allow matching of synthetic control and treated firm at the level of broad import bundles.

control firms for treated and placebo firms.[41] Larger donor pools and the volatility of the outcome variables substantially slowed the underlying SCM fitting procedure.

We have very large potential donor pools made up of all the Serbian firms that imported between 2010 and 2013. Consistent with other applications of the SCM we limit the donor pools to firms that share some similarities. In particular, firms that are potential donors and/or placebos are limited to being within the same decile of 2010 total import value as the treated firm. The treated firms for which there are continuous imports are all within the 4th decile or higher in terms of their 2010 import value.

As noted in Section 3.2 we do not draw placebo firms from the donor pool used to construct the synthetic control firm for the estimated treatment effect of the IHC. Rather, for each treated firm we first randomly allocate untreated firms in the 2010 import value decile to placebo or donor status so that placebo firms are not eligible to be components of the synthetic control firm for the treated firm (nor can they be components of the synthetic control firm for other placebos). The synthetic control units for both treated and placebo firms are constructed by drawing from the same (large) pool of potential donors. For our benchmark (and preferred) specification, with three pre-treatment values of the outcome variable as conditioning variables, as well as for the majority of other specifications a potential donor pool of 125 firms was large enough to allow synthetic control firms to be constructed for treated and placebo firms for all outcome variables.[42]

5. Results

5.1 Difference-in-Differences Results

Before presenting our main results from the pooled SCM, as a benchmark we provide results from DiD regressions to estimate the effect of IHC adoption on the five firm-level outcomes discussed in Section 4.1.[43] The DiD estimator compares the differences in an outcome variable

---

[41] The restriction that the w* weights all be positive and sum to one forces the synthetic control firm to be a convex combination of the firms that are used to construct it. Smaller donor pools may not contain a sufficient variety of firms to allow the joint fitting of time series outcomes and a large number of firm characteristics.

[42] In one of the robustness exercises (log import value with the average lag of log imports included in the characteristic variables), we required a larger set of placebo firms to insure the construction of synthetic control firms for two of the treated firms and their respective placebos. Since 125 potential donors were not sufficient, we expanded the size of the donor pool to 200. Because our inference relies on rank order of the treated firm, relative to its placebo firms, it is not directly affected by varying the donor pool size so long as we give the treated firm and its placebo access to the same donor pool.

[43] While we use logarithms and moving averages to smooth the dependent variables, which as discussed in Section 4, is important for the SCM estimation, the qualitative DiD results we present here follow through if we use the log of monthly variables rather than the three-month moving average of that log.

across treated firms that adopt the IHC program and control firms that do not adopt the IHC program in the periods before and after the firms adopt the IHC program. Our benchmark DiD regression includes a treatment dummy defined for each treated firm as 1 for any month after the firm has begun to clear imports using the IHC program and 0 otherwise and defined for each control firm as 0 in every month, as well as firm and month fixed effects. The DiD regression is a simplified version of Eq. (2) where $\alpha_{jt} = \alpha$ (i.e., there is a single average treatment effect), λ_t is constant over time (i.e., firm fixed effects are included), and the vector of conditioning variables is ignored ($\boldsymbol{\theta}_t$=0). Our DiD regressions are estimated based on a sample that includes all 21 treated firms and a set of control firms (that did not adopt the IHC program), from within the same deciles (4-10) of total 2010 import value as the treated firms. This choice of control firms is useful for comparison to the SCM results. The DiD regression results are reported in Table 6. The estimates show a significant negative effect of the IHC program on customs clearance times at various points in the distribution but no effect on inspection rates or import value.

Table 6. Difference-in-Differences Regressions, Pooling Across Treated Firms

	Median clearance time	Clearance time 75th pct.	Clearance time 90th pct.	Inspection rate	Import value
Treatment dummy	-0.223***	-0.508***	-0.549***	-0.007	0.035
	(0.043)	(0.049)	(0.056)	(0.006)	(0.034)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
Month fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	50,738	50,738	50,738	50,738	50,738
R-squared	0.777	0.759	0.734	0.741	0.857

Notes: Standard errors in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% confidence levels, respectively. The dependent variables are 3-month moving averages of the variables in logs (with the exception of inspection rate that is in levels). The sample is limited to the treated firms and the control firms included in deciles 4-10 of firm total import value in 2010.

The DiD estimator, however, relies on strong identifying assumptions. It requires that, in the absence of treatment, the outcome variables for treated and control firms would have experienced similar time trends. In order to test this assumption, we consider data for the 21 treated firms and control firms in the pre-treatment period (defined to be all months prior to July 2011, when the first firm adopted the IHC program) and estimate a regression of each of the five outcome variables on a common time trend and the interaction of the trend with a dummy for future treatment indicating that the firm will adopt the IHC program within our sample period. As in Table 6, we estimate each regression for a sample including all treated firms and only control firms

in the same size decile as the treated firms. The significance of the coefficient on the interaction term thus serves as our test statistic for the validity of the common trends assumption. We report these results in Table 7. The results indicate that the common trends assumption is violated for three of the five outcome variables of interest. We therefore view the estimates in Table 6 as invalid.

Table 7. Test of Common Trends Assumption

	Median clearance time	Clearance time 75th pct.	Clearance time 90th pct.	Inspection rate	Import value
Trend	-0.014***	-0.015***	-0.012***	-0.001***	0.015***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Trend*future treatment dummy	0.024***	0.021***	0.006	0.001	0.018***
	(0.006)	(0.007)	(0.008)	(0.001)	(0.005)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
Month fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	17,648	17,648	17,648	17,648	17,648
R-squared	0.844	0.832	0.806	0.767	0.885

Notes: Standard errors in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% confidence levels, respectively. The dependent variables are 3-month moving averages of the variables in logs (with the exception of inspection rate, which is in levels). The sample is limited to the treated firms and the control firms included in deciles 4-10 of firm total import value in 2010 and covers only months prior to the initial month in which the IHC program was used (July 2011).

One potential problem with these results is that we are pooling across different firms that adopt the IHC program and across firms in the control group. As discussed in Section 3, the IHC program may have heterogeneous effects across different types of firms. Also, firms of different sizes may have different underlying trends in the outcome variables. In order to check this possibility, we estimate the DiD regression separately for each treated firm, using as a control group firms in the same size decile as the treated firm. Such regressions have fairly weak statistical power because they include a single treated firm, but they are nonetheless useful for showing the weakness of a DiD estimator for this application.

Table 8 offers a summary of the coefficient estimates from 105 DiD regressions and from 105 common trend testing regressions (105 = 21 treated firms x 5 outcome variables). The estimated means of the coefficients on the IHC treatment dummy are of the expected signs. There are relatively high numbers of statistically significant coefficients for clearance times and for import value, considerably more than would be expected from random variation. The common

trends assumption is violated 19 times in the 105 regressions, and 10 of the 19 violations occur in DiD regressions whose coefficient estimates would have otherwise led us to conclude there was a statistically significant effect of the IHC program. This evidence, together with the pooled estimates from Table 7 suggest that the DiD estimator is not suitable for inference for any of our outcome variables of interest in these data because the common trends assumption is too frequently violated.[44] This motivates our use of the SCM below. An additional lesson from Table 8 is the evidence of potential heterogeneity in the estimated effects of the IHC program across treated firms. The DiD regressions for import value in particular generate a relatively large number of significantly positive and significantly negative effects.

Table 8. Results Summary from Firm-by-Firm DiD Regressions and Common Trend Testing Regressions

Dependent variable	Median estimated coefficient on treatment variable	Mean estimated coefficient on treatment variable	Number of significant coefficients > 0	Number of significant coefficients < 0	Number of violations of common trends assumption	Number of significant treatment coefficients with common trends violation
Median clearance time	-0.259	-0.321	1	7	5	2
Clearance time 75th pct	-0.441	-0.571	0	10	4	3
Clearance time 90th pct	-0.365	-0.556	1	8	5	3
Inspection rate	-0.004	-0.007	0	1	2	1
Import value	0.083	0.048	6	2	3	1

Notes: The table summarizes regression results from 105 DiD regressions and 105 regressions that test for common trends. Each of the 21 firms adopting the IHC program appears in one treatment regression and in one common trends testing regression, and the control firms are those within the same size decile as the treated firm. For each dependent variable, the median and mean of the estimated treatment effects across the 21 firms are reported, as well as the number of DiD regressions (out of 21) with significant positive and negative coefficients at the 5% confidence level, respectively, for the treatment dummy. The final two columns show, respectively, the number of regressions that indicate a statistically significant violation of the common trends assumption at the 5% confidence level and the number of cases in which a statistically significant coefficient coincided with a violation of the common pre-treatment trends.

5.2 Benchmark Synthetic Control Method Results

This section presents the SCM estimates for the impact of the IHC program on importing firms, considering five outcomes of interest and addressing the following questions. First, did the IHC program reduce clearance times for the firms that adopted it? We evaluate this question at

---

[44] Given a confidence level of 5%, one might expect only one failure of the common trends assumption to arise in 21 regressions, or roughly five in 105 regressions.

three points in the clearance time distribution: the median, the 75th percentile and the 90th percentile. Second, did the program reduce inspection rates for IHC firms? Third, did the program lead firms to increase the value of their monthly imports?

As mentioned in Section 2.2, 21 firms adopted the IHC program early enough in the 2010-2013 sample period so that we have sufficient months of data prior to and subsequent to that adoption. The first step in applying the SCM to the IHC program evaluation is to construct for each of those 21 firms a synthetic control firm based on the donor pool, that as mentioned in Section 4.2, we restrict to a randomly selected set of 125 importing firms in the same size decile as the treated firm.[45] For each treated firm other than the first firm to adopt IHC the SCM estimation is based on a pre-treatment period of 19 months. For the first IHC adopter that period is of 17 months due to data availability constraints. In the algorithm from which weights are derived to construct the synthetic control firm (as a convex combination of firms in the donor pool) for each treated firm we use the conditioning variables described in Section 4.1: the share of imports originating in the EU, the share of imports in each sector, and the share of import declarations receiving special customs treatment, all averaged over the entire pre-treatment period as well as three pre-treatment values of the outcome variable in the first, 10th, and 18th months in the pre-treatment period. As an illustration of the synthetic control method estimator, we describe in detail in Appendix C the analysis based on graphical representations for one specific IHC-adopting firm.

In order to pool the analysis across IHC firms for each outcome variable we construct the $\hat{\beta}_f$ statistic defined in Eq. (7) by averaging the $\hat{\alpha}_{ft}$'s across the first 6 post-treatment months for each of the IHC firms and for their associated placebos. Focusing on the first outcome variable, log median clearance time, we denote this statistic as $\widehat{\beta\_hrs}_f$. As described in Section 3.2, we then calculate the rank of the treated firm's $\widehat{\beta\_hrs}_f$ among the distribution of $\widehat{\beta\_hrs}^P$ generated by its placebos and the percentile rank statistic for each IHC firm. Table 9 shows the estimated impact of the IHC program on log median clearance time for each firm, with firms ordered by the timing of their initial use of the IHC program. In addition to the $\widehat{\beta\_hrs}_f$ statistic, the table reports for each IHC firm the number of donor firms used to construct its synthetic control firm, the rank

---

[45] As discussed in Section 4.3, in two cases where the SCM procedure did not construct a synthetic control firm for either the treated firm or one of the placebo firms using 125 donor firms we expanded the pool of donors to 200 so both the treated and placebo firm until a synthetic control firm could be successfully constructed. Because our inference relies on the impact on the treated firm, relative to its placebos, a consistent treatment across treated and placebo firms is enough to support our inferences.

of its $\widehat{\beta\_hrs}_f$ statistic among those of its placebos, and the associated percentile rank statistic.[46] At the bottom, the table also shows the mean and median values of the $\widehat{\beta\_hrs}_f$ statistic across the 21 IHC firms and the mean of the percentile rank statistic across the IHC firms, the statistic $\bar{p}$. The average estimated (log) median clearance time reduction caused by the IHC program is, 0.19, or an approximate 19 percent reduction in firms' median clearance times. This estimate is affected by one very large negative outlier (the estimate for firm #19). The median estimated effect across the treated firms is smaller, at 0.073. In order to test for a significant impact of the IHC program on clearance times as in Dube and Zipperer we calculate the rank of each firm's $\widehat{\beta\_hrs}_f$ statistic, relative to that of its placebo firms, and the associated percentile rank statistic. The mean of these percentile rank statistics is 0.434, which is not low enough to indicate a significant impact of the IHC program on log median clearance times.[47] Similar tables with benchmark results for each of the other four variables are reported in Appendix B.

In order to discuss in detail the significance of the SCM results, point estimates of the impact of the IHC program can be expressed in terms of the median treatment effect, the mean treatment effect (both of which were already shown in Table 9 for the median clearance time outcome), or through the Hodges-Lehmann (henceforth HL) estimate. Dube and Zipperer (2015) propose the HL estimate as the most robust to outliers, and we use it as our preferred point estimate of the mean impact of the IHC program. The HL confidence interval also provides a sense of the range of effect sizes that are plausible, given our estimates.

The various estimates for our benchmark specification for the five firm-level outcome variables are reported in Table 10. The main lessons are that the variability of the estimates makes it difficult to attribute causal effects of the IHC program at standard levels of statistical significance, with the exception of changes at the 75th percentile of the clearance time distribution. Perhaps more important for policy officials interested in the impact of trade facilitation reforms are the confidence intervals, which offer bounds on the size of the treatment effect. For median clearance times, for clearance times at the 90th percentile and for import value, the intervals contain zero but they also contain large changes in the direction one would expect (that is, lower clearance times and higher firm-level imports).

---

[46] In examining the ranks for the 21 IHC firms relative to their placebos in Table 9, note that low ranks imply short clearance times.
[47] The two-sided test of the null hypothesis of no impact can be rejected at a 5 percent significance level if the mean percentile rank statistic falls below 0.377 or above 0.623, which are the critical values based on Appendix table A1 in Dube and Zipperer (2015). The critical values for a 1 percent significance level are 0.339 and 0.660.

Table 9. Synthetic Control Method Detailed Results for All IHC Firms - Median Clearance Time

IHC firm #	6-month average treatment effect: change in log hours ($\widehat{\beta\_hrs}_f$)	Number of donor firms used to construct synthetic control firm	Rank of IHC firm relative to 25 placebo firms	Percentile rank statistic (rank/26)
1	-0.049	10	10	0.385
2	-0.240	14	5	0.192
3	-0.060	8	8	0.308
4	-0.822	11	2	0.077
5	-0.308	13	3	0.115
6	0.496	6	22	0.846
7	-0.094	11	12	0.462
8	-0.195	8	11	0.423
9	-0.783	8	3	0.115
10	0.122	10	17	0.654
11	-0.104	13	9	0.346
12	-0.510	8	3	0.115
13	0.183	4	15	0.577
14	0.734	5	24	0.923
15	0.281	100	21	0.808
16	-0.005	5	16	0.615
17	0.445	3	23	0.885
18	-0.073	3	10	0.385
19	-2.779	3	1	0.038
20	-0.036	3	13	0.500
21	-0.246	7	9	0.346
Mean % change in hours	-0.192		Mean percentile rank statistic ($\bar{p}$)	0.434
Median % change in hours	-0.073			

Notes: Synthetic control firms were constructed with the benchmark specification that includes three pre-treatment values of the outcome variable as part of the set of conditioning variables. The summary statistics at the bottom left of the table are obtained as the mean and the median of the treatment effect across the 21 firms. The statistic at the bottom right of the table $\bar{p}$ given 21 firms has 5% critical values of 0.377 and 0.623 (based on Appendix table A1 in Dube and Zipperer (2015)).

Table 10. Benchmark Estimates of the Impact of the IHC Program on Outcome Variables

	Median treatment effect	Mean treatment effect	Mean percentile rank statistic ($\bar{p}$)	Hodges Lehmann	
				Estimate	95% confidence interval
Median clearance time	-0.073	-0.192	0.434	-0.064	(-0.216, 0.067)
Clearance time: 75th pct	-0.401	-0.434	0.310***	-0.404	(-0.618, -0.167)
Clearance time: 90th pct	-0.362	-0.410	0.377*	-0.261	(-0.626, 0.004)
Inspection rate	-0.007	-0.008	0.484	-0.002	(-0.021, 0.015)
Import Value	0.106	0.043	0.588	0.117	(-0.049, 0.253)

Note: * and *** indicate significance at the 10% and 1% confidence levels, respectively.

While the confidence intervals in Table 10 look wide in percentage terms, for clearance times at least the large variability in percentage terms applies to a very small base. In order to better understand the practical importance of our estimates we conduct some back-of-the-envelope calculations and report them in Table 11. Table 3 showed that the distribution of clearance times in 2010 for firms that would become IHC firms by 2013. Application of the HL estimates to the median 2010 clearance time for firms that would become IHC firms (0.84 hours, or 50 mins) implies an estimate of the change in median clearance time of 3 minutes, with the 95% confidence interval stretching from a reduction of 11 minutes to an increase of 3 minutes. At the 75th percentile of clearance times, the HL estimate suggests a 42 minute reduction in clearance times (relative to a 103 minute benchmark). The 95% confidence interval puts the true reduction somewhere between 17 and 64 minutes. At the 90th percentile of the clearance time distribution, the HL estimate suggests a 57 minute reduction, and the 95 confidence interval is bounded between a 137 minute decrease and a 1 minute increase.

Table 11. Implied Effect of the IHC Program on Clearance Times (In Minutes)

	2010 benchmark clearance time	Implied reduction, HL estimate	Implied reduction 95% HL confidence interval
Median	50	-3	(-11, 3)
75th percentile	103	-42	(-64, -17)
90th percentile	218	-57	(-137, 1)

Notes: Changes in clearance time calculates by applying percentage change estimated in Table 10 to relevant locations in the distribution of 2010 clearance times for firms that would become IHC firms by the end of 2013. `HL' stands for Hodges-Lehmann estimate.

6. Synthetic Control Method Robustness Results

We take three approaches to checking the robustness of the SCM estimates. First, we restrict the sample to the first 15 adopters of the IHC program, which allows us to measure impact

over a 9-month rather than a 6-month post-treatment window. Second, we take two alternative approaches to constructing the synthetic control firms, using alternative approaches to incorporating lagged outcome variables into the set of conditioning variables $\mathbf{X_j}$. Third, we return to the benchmark specification, but check the robustness of the pooled estimates to the exclusion of poorly fitting placebos. A summary of what follows is that the main results from the benchmark specification hold up across all these robustness checks: 1) there is strong evidence that clearance times fell at the 75th percentile of the clearance time distribution; 2) the evidence that clearance times fell at the 90th percentile is mixed, but suggests that a reduction is quite likely; 3) there is no statistically significant change in median clearance times or firm-level import value, but there are wide confidence intervals that also contain the possibility of large effects in the predicted direction, and 4) very little evidence points to a significant change in inspection rates for IHC firms.

We report the results of the first two robustness checks in Table 12. For each outcome variable, the first row of Table 12 reports results when we consider only the first 15 adopting firms (those that began using the IHC program before April 2013). Using only the earliest adopters allows us to use a longer time window of 9 months to assess program impact.[48] The tradeoff is that the smaller sample reduces statistical power and increases the width of confidence interval. For all five outcome variables the results for the 9-month window are in line with those for the 6-month window reported in Table 9.[49]

Following Dube and Zipperer, the next set of robustness checks uses different conditioning variables $\mathbf{X_j}$ in the construction of the synthetic control firm. Specifically, the approaches use information from lagged outcome variables differently. One approach includes all pre-treatment outcome variables in the set of conditioning variables. This produces a synthetic control firm that matches the pre-treatment dynamics of the treated firm very well, but can mean that other conditioning variables (e.g. the shares of various commodities in the import bundle) do not fit as well. The other approach includes only the average of the pre-treatment outcome variables, none of their specific values are used in the set of conditioning variables. This approach produces a worse fit of pre-treatment dynamics, but typically generates very good matches with respect to the

---

[48] In principle the variable that would change most slowly in response to the IHC program is import value. It may take some time for firms to change the level of their imports after the program reduces trade costs. Data limitations preclude us from estimating over longer post-treatment windows, and although the 9-month window is not so much longer it allows at least some scope for impacts to emerge over time.

[49] In unreported results we also considered a 3-month window at the end of the 9-month window. If an effect were to emerge over time this would limit bias that comes with averaging over periods of reduced impact. Our results were qualitatively and quantitatively similar to results we report for the benchmark and the robustness checks.

conditioning variables. Table 12 shows that both approaches return qualitatively and quantitatively similar estimates of the causal impact of the IHC program across all five outcome variables, relative to Table 9.

Table 12. Estimates of the Impact of the IHC Program on Outcome Variables – Robustness

				Hodges Lehmann	
	Median treatment effect	Mean treatment effect	Mean percentile rank statistic ($\bar{p}$)	Estimate	95% confidence interval
(log) Median clearance time					
9-months post treatment	-0.097	-0.090	0.428	-0.077	(-0.240, 0.083)
Match all pre-treatment outcomes	-0.098	-0.153	0.436	-0.074	(-0.213, 0.082)
Match average pre-treatment outcomes	-0.177	-0.247	0.401	-0.161	(-0.346, 0.025)
(log) Clearance time – 75th percentile					
9-months post treatment	-0.522	-0.440	0.308***	-0.687	(-0.687, -0.065)
Match all pre-treatment outcomes	-0.436	-0.467	0.305***	-0.419	(-0.659, -0.178)
Match average pre-treatment outcomes	-0.407	-0.514	0.317***	-0.379	(-0.725, -0.108)
(log) Clearance time – 90th percentile					
9-months post treatment	-0.207	-0.408	0.379	-0.205	(-0.649, 0.075)
Match all pre-treatment outcomes	-0.497	-0.554	0.320***	-0.490	(-0.834, -0.148)
Match average pre-treatment outcomes	-0.497	-0.625	0.344**	-0.516	(-0.877, -0.109)
Inspection rate					
9-months post treatment	-0.008	-0.015	0.420	-0.006	(-0.023, 0.013)
Match all pre-treatment outcomes	-0.014	-0.016	0.421	-0.011	(-0.027, 0.006)
Match average pre-treatment outcomes	-0.011	-0.010	0.440	-0.006	(-0.025, 0.012)
(log) Import value					
9-months post treatment	0.076	0.009	0.562	0.082	(-0.090, 0.255)
Match all pre-treatment outcomes	0.044	-0.032	0.529	0.028	(-0.146, 0.171)
Match average pre-treatment outcomes	0.054	0.036	0.562	0.097	(-0.084, 0.245)

Notes: ** and *** indicate significance at the 5% and 1% confidence levels, respectively. The 95% critical values for the mean percentile rank statistic in the case of 15 firms (used with the 9-month treatment period) are 0.354 and 0.646. The 95% critical values for the other exercises (with 21 firms) are 0.377 and 0.623.

Our preferred SCM estimates, which we call the benchmark case, include three pre-treatment outcome variables in the conditioning set. The choice of these estimates as the benchmark case is supported by a data-driven metric proposed by Dube and Zipperer. For each estimator we calculate the root mean squared prediction error (RMSPE) for both the pre-treatment period and for the six-month post-treatment windows among the placebo firms. Since the placebos

lack any treatment by the IHC program, the post-treatment RMSPE for the placebo firms is a good indicator of how well SCM forecasts the outcome variable. The results are reported in Table 13. Dube and Zipperer suggest that the preferred set of conditioning variables should be those that produce the smallest post-treatment RMSPE. By this metric, the benchmark set of conditioning variables is the set that performs best for each of the five outcome variables.

Table 13. Forecast Accuracy among Placebo Firms

	RMSPE for placebo firms		
	Pre-treatment	Post-treatment	95% CI for population median of forecast error
(log) Median Clearance Time			
Benchmark	0.356	0.629	(-0.008, 0.062)
Match all pre-treatment outcomes	0.271	0.631	(0.002, 0.080)
Match average pre-treatment outcomes	0.420	0.651	(-0.028, 0.017)
(log) Clearance time - 75th percentile			
Benchmark	0.405	0.740	(-0.049, 0.068)
Match all pre-treatment outcomes	0.291	0.772	(-0.040, 0.073)
Match average pre-treatment outcomes	0.485	0.793	(-0.137, -0.026)
(log) Clearance time - 90th percentile			
Benchmark	0.482	0.917	(-0.076, 0.057)
Match all pre-treatment outcomes	0.328	0.961	(-0.089, 0.043)
Match average pre-treatment outcomes	0.565	0.966	(-0.150, 0.008)
Inspection rates			
Benchmark	0.053	0.090	(-0.009, 0.001)
Match all pre-treatment outcomes	0.040	0.094	(-0.011, 0.004)
Match average pre-treatment outcomes	0.062	0.098	(-0.010, 0.001)
(log) Import value			
Benchmark	0.340	0.564	(-0.042, 0.033)
Match all pre-treatment outcomes	0.247	0.584	(0.007, 0.106)
Match average pre-treatment outcomes	0.419	0.587	(-0.040, 0.001)

Note: For each set of conditioning variables, the table reports root mean square prediction errors for pre- and post-treatment periods (columns 2 and 3). The table also reports, for each set of conditioning variables the confidence interval for the population median of the forecast error in the 6th month of the post-treatment period (column 4).

The ratio of post-treatment RMSPEs to pre-treatment RMSPEs is relatively high, relative to Dube and Zipperer's minimum wage experiments. Our outcome variable series are more volatile, and the forecast errors are larger. In order to check for possible biases in the SCM estimator we collect the realized error terms in the 6th month of the post-treatment period for the 525 placebo firms for each outcome variable and SCM estimator. Column 4 of Table 13 reports the 95% confidence interval for the population median of the forecast errors.[50] While four of the 15 cases do not contain zero, these cases are all among the robustness checks. For the benchmark SCM estimator, the confidence intervals for all five variables contain zero. We take this as additional evidence that the benchmark model is the preferred SCM estimator for these data.

Dube and Zipperer propose an additional approach to checking robustness, excluding poorly fitting placebos from the analysis. Poorly fitting placebos might bias the pooled SCM estimator against finding impact if they artificially expand the implied distribution of random outcomes. The robustness check involves discarding the placebos with pre-treatment RMSPEs that are large multiples of that of the treated firm. They vary this multiple (from 20 to 2) in their robustness check, as do we. We report the results from this exercise in Table 14 which are qualitatively and quantitatively in line with those of the benchmark SCM estimator. There is a strong and statistically significant negative effect of the IHC program on clearance times at the 75th percentile of the clearance time distribution. The evidence for reductions at the 90th percentile is mixed, but strongly suggestive of impact. The 95 percent confidence intervals either barely include zero or do not include it, but also include large negative effects on clearance times at the 90th percentile. Impacts on median clearance times and import value are not statistically significant, but the confidence intervals also suggest the possibility of large impacts in the expected direction (negative for median time and positive for import value).

---

[50] We focus on the population median since the population mean is sensitive to outliers, and the rest of our analysis has sought to reduce that sensitivity.

<p style="text-align:center">Table 14. Robustness to Dropping Poorly Fitting Placebos</p>

RMSPE Ratio	Number of placebo firms used	Fraction of all placebo firms used	Mean percentile rank statistic ($\bar{p}$)	Estimate	Hodges Lehmann 95% confidence interval
			(log) Median Clearance Time		
.	525	1.00	0.434	-0.065	(-0.216, 0.067)
20	486	0.93	0.427	-0.068	(-0.216, 0.054)
10	449	0.86	0.428	-0.064	(-0.215, 0.056)
5	410	0.78	0.422	-0.064	(-0.215, 0.052)
2	318	0.61	0.409	-0.081	(-0.243, 0.033)
			(log) Clearance time - 75th percentile		
.	525	1.00	0.309***	-0.404	(-0.618, -0.167)
20	479	0.91	0.310***	-0.404	(-0.611, -0.157)
10	443	0.84	0.310***	-0.382	(-0.597, -0.180)
5	406	0.77	0.310***	-0.381	(-0.605, -0.198)
2	329	0.63	0.334***	-0.404	(-0.677, -0.105)
			(log) Clearance time - 90th percentile		
.	525	1.00	0.377*	-0.261	(-0.626, 0.004)
20	483	0.92	0.372**	-0.259	(-0.613, -0.011)
10	451	0.86	0.374**	-0.234	(-0.583, -0.004)
5	409	0.78	0.369**	-0.238	(-0.580, -0.021)
2	325	0.62	0.383*	-0.194	(-0.469, 0.008)
			Inspection rates		
.	525	1.00	0.484	-0.002	(-0.021, 0.015)
20	474	0.90	0.463	-0.005	(-0.023, 0.012)
10	428	0.82	0.464	-0.004	(-0.022, 0.012)
5	361	0.69	0.463	-0.004	(-0.022, 0.012)
2	236	0.45	0.496	-0.001	(-0.018, 0.014)
			(log) Import value		
.	525	1.00	0.588	0.110	(-0.049, 0.253)
20	494	0.94	0.589	0.106	(-0.049, 0.246)
10	466	0.89	0.590	0.102	(-0.048, 0.240)
5	424	0.81	0.585	0.089	(-0.048, 0.214)
2	338	0.64	0.620*	0.127	(-0.005, 0.245)

Notes: *, **, and *** indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Placebos are assessed by the match quality of the synthetic control firm constructed by the benchmark SCM model. The RMSPE ratio is the ratio of pre-treatment root mean square prediction error for the placebo relative to that of the treated firm. Rows show results for sub-samples where placebos are limited to those whose RMSPE ratios are less than X, where X = . indicates the full sample.

7. Testing for Heterogeneity

The tests presented so far evaluate the null hypothesis of no impact of the IHC program on the outcome variables of interest for all treated firms. Formally, the null hypothesis could be rejected if there is a common impact on all firms, or if there is large enough impact on a subset of firms (even just one) without any impact on the remaining firms. To better inform these alternatives, we apply tests for heterogeneous impacts of the IHC program across treated firms.

As noted in Section 3.4, in cases where there is evidence of a significant treatment effect, a test on how the distribution of percentile rank statistics based on mean-adjusted treatment effects (below referred to as *mean-adjusted percentile rank statistics*) compares to a uniform distribution is useful for understanding whether or not the impacts are likely to be common across treated firms. When there is no evidence of a significant treatment effect, a test on how the distribution of the unadjusted (original) percentile rank statistics compares to a uniform distribution is useful for understanding whether or not there might be offsetting effects on either side of the mean treatment effect.

Figure 1 presents results from the Kolmogorov-Smirnov and the Anderson-Darling heterogeneity tests for four of our five outcome variables using the estimates from our benchmark specification (described in Section 5).[51] For each outcome variable, we provide two graphs: 1) in the left-hand column a representation of the distribution of the unadjusted percentile rank statistics generated by SCM, and 2) in the right-hand column a representation of the distribution of the mean-adjusted percentile rank statistics. The first graph offers additional evidence on the null hypothesis of no impact. The second graph is most useful for testing the commonality of impact, given that an impact exists.

The tests for heterogeneity in the unadjusted results are largely consistent with the results of the earlier tests. Unadjusted percentile rank statistics for the 75th percentile of the clearance time distribution are not distributed uniformly, as they would be under the null hypothesis of no impact. There is some evidence of departures from uniformity for unadjusted percentile rank statistics for the 90th percentile of the clearance time distribution and for import value, but these are not statistically significant. The primary lesson from the left-hand column is that our earlier tests did not appear to miss any statistically significant effects that arose because the mean percentile rank statistic obscured heterogeneous treatment effects across firms.

The right-hand column is more interesting, because it provides evidence on the question of whether the treatment effect identified by previous tests on the 75th percentile of clearance time distribution was heterogeneous across firms or common to all firms. If mean-adjusted percentile rank statistics look to be uniformly distributed, we should accept the null hypothesis that the treatment impact is common across all firms. This appears to be the case. We cannot reject the

<hr />

[51] For purposes of visual clarity we exclude graphs for the inspection rate because there is little evidence in any of our treatments for a causal effect of the IHC program on the inspection rate. There is no evidence of firm heterogeneity in the causal effects of inspection rates, nor is there evidence of heterogeneity remaining when we adjust by the quite modest mean estimated treatment effect.

null of a homogeneous impact of the IHC program on the 75th percentile of the clearance time distribution across treated firms.

Figure 1. Heterogeneity Tests for Benchmark Synthetic Control Method Estimates
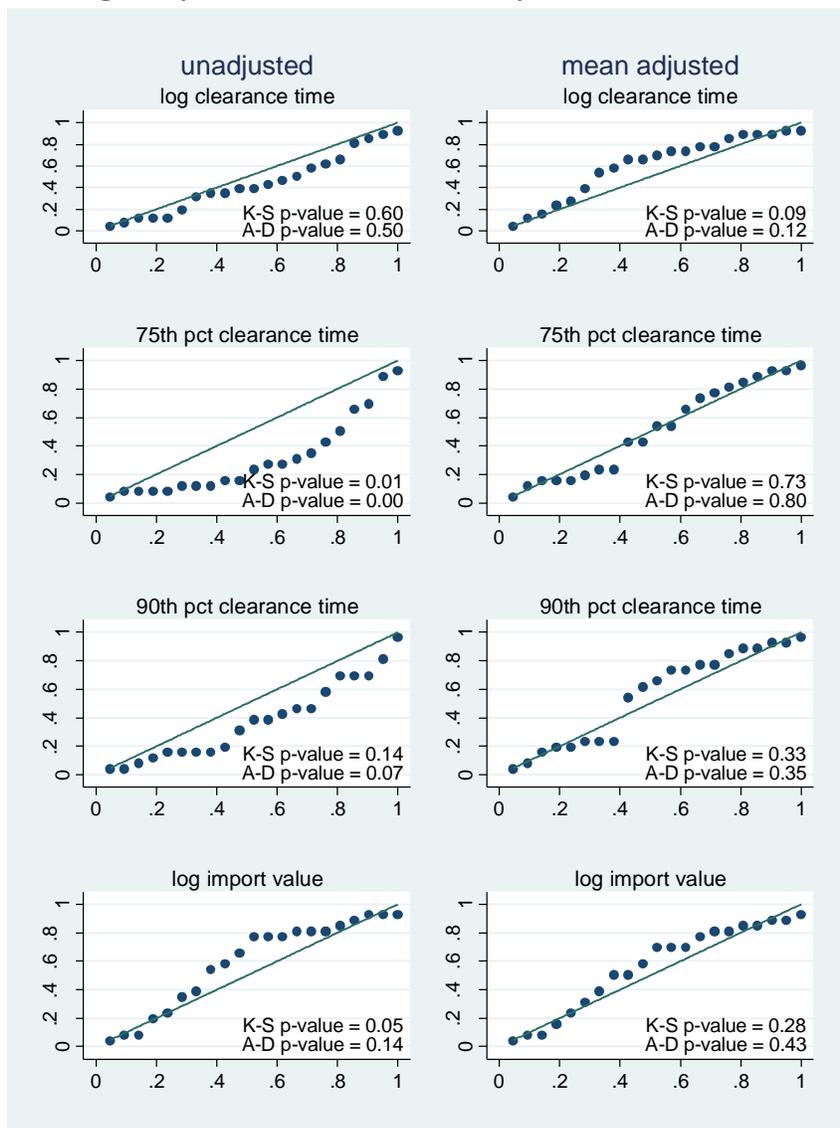
Notes: The figure shows the distribution of percentile rank statistics, p_f, for each of four outcome variables, relative to a cumulative distribution function of the uniform distribution function. The first column shows distributions of p_f from the benchmark SCM results. The second column shows distributions of p_f recalculated after the estimated treatment effects are adjusted by the mean estimated treatment effect ($\bar{\beta}$). Formal tests for heterogeneity in each case are either Kolmogorov-Smirnov (K-S) or Anderson-Darling (A-D). The p-values for each test are reported in each graph.

8. Conclusion

Each year hundreds of millions of dollars in development aid are spent supporting trade facilitation reforms. Yet few of these reforms have been subject to formal scrutiny via impact evaluation. Procedural and other reforms by border management agencies are undertaken in

complex environments that include differential treatment for heterogeneous agents, complex dynamics, plausibly heterogeneous treatment effects across units, and real-time operational demands that typically rule out randomization as an evaluation design. All of these frustrate conventional approaches to program evaluation. In this paper we apply a recently developed approach to pooling across estimates from the synthetic control method. We apply this technique to a recent customs reform in Serbia that allowed qualified firms to clear their import shipments through customs at their own warehouses rather than at the customs office.

Our pooled synthetic control estimates indicate that as a result of the IHC program clearance times for treated firms fell significantly at the 75th percentile and (with weaker significance) at the 90th percentile. While there are relatively large confidence intervals for the effects when measured as percentage changes, since the reductions are relative to small bases the treatment effects and the uncertainty about them are much more modest in level terms. In subsequent heterogeneity tests we are unable to reject the null hypothesis that the treatment effect of the IHC program on these variables is common across treated firms. The high degree of variability in the treatment effects means that we are unable identify statistically significant impacts of the IHC program on median clearance times or import value. The confidence intervals for these outcomes do suggest that bad outcomes as a result of the IHC program are unlikely and some sizable treatment effects in the expected direction are possible. We find little evidence for any effect of the program on inspection rates.

In broad terms the relatively weak effects that we estimate here should be understood as linked to selection into the program. The firms that adopted the IHC program were already very well treated by Serbian customs before the program began. While the program offers additional benefits to these firms, it does not appear to fundamentally change much about clearance times or inspection rates. Hence, the little evidence of impact on import value that we find is not surprising. If firms that were not already so well treated by customs could enter the IHC program it may well be that larger impacts would emerge. Another reason for our weak effects is that the outcome variables we study are noisy, thus there are rather large standard errors around the estimated treatment effects. The confidence intervals we report usually contain zero but do allow for the possibility of substantial treatment effects in the direction that would be expected.

We have not studied the systemic benefits that might flow from the IHC program. Our data suggests that in 2013 the program allowed approximately 2.6 percent of import value to be

cleared outside of the customs office. Presumably these were low-risk shipments, and clearing them offsite offers operational benefits to the customs agency, which can focus its attention on the remaining, higher risk shipments. Moving IHC shipments away from the customs office may also offer small decongestion benefits that flow to other firms. Our estimation strategy assumes the decongestion benefits are non-existent but they may exist, and they are certainly plausible if the program were to expand. It is also possible that the program would have statistically significant impacts on firms that were less able to adopt the program. Support for this claim would require further research.

While a better understanding of the effect of the IHC program is useful, the larger purpose of this study is to introduce pooled synthetic control methods into the trade facilitation literature. Customs reforms are particularly difficult reforms to evaluate because the trade outcomes they seek to influence exhibit complex dynamics and are often heterogeneous across units. We believe the pooled synthetic control method offers an opportunity to extend rigorous evaluation methods to a body of reforms that are otherwise quite difficult to evaluate credibly.

References

Abadie, A. and J., Gardeazabal (2003). "The Economic Costs of Conflict: a Case Study of the Basque Country", *American Economic Review* 93:113-132.

Abadie, A., Diamond, A., and J., Hainmuller (2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association* 105: 493-505.

Atkin, D., Khandelwal, A., and A. Osman (2014). "Exporting and Firm Performance: Evidence from a Randomized Trial," NBER Working Paper 20690.

Billmeier, B. and T. Nancinni (2013). "Assessing Economic Liberalization Episodes: A Synthetic Control Approach," *Review of Economics and Statistics* 95: 983-1001.

Cadot, O., Fernandes, A., Gourdon, J. and A. Mattoo (2015). "Are Export Support Programs Effective? Evidence from Tunisia," *Journal of International Economics* 97: 310-324.

Carballo, J., Schaur, G. and C. Volpe Martincus (2016). "Trust No One? Security and International Trade," University of Tennessee mimeo.

Castillo, V., Garone, L., Maffioli, A., and L. Salazar (2015). "Tourism Policy, a Big Push to Employment? Evidence from a Multiple Synthetic Control Approach," *IDB Working Paper Series*, 572.

Clark, D., Schaur, G. and V. Kozlova (2013). "Supply Chain Uncertainty as a Trade Barrier," University of Tennessee mimeo.

Cunningham, S. and M. Shah (2014). "Decriminalizing Indoor Prostitution: Implications for Sexual Violence and Public Health" *NBER Working Paper* # 20281.

Deaton, A. (2010). "Instruments, Randomization and Learning about Development," *Journal of Economic Literature* 48: 424-455.

Dube, Arindrajit and Ben Zipperer (2015) "Pooling Multiple Case Studies using Synthetic Controls: An Application to Minimum Wage Policies," *IZA Discussion paper* 8944.

Fernandes, A., Hillberry, R. and A. Mendoza-Alcantara (2015). "Trade Effects of Customs Reform," *World Bank Policy Research Working Paper* # 7210.

Heckman, J., Ichimura, H. and P. Todd (1997). "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies* 64: 605-654.

Heilmann, K. (2016). "Does Political Conflict Hurt Trade? Evidence from Consumer Boycotts," *Journal of International Economics* 99: 179-191.

Hirano, K. Imbens, G., and G. Ridder (2003). "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica* 71: 1161-1189.

Hodges Jr., J.L. and E.L. Lehmann (1963). "Estimates of Location Based on Rank Tests," *Annals of Mathematical Statistics* 34: 598-611.

Hummels, D., and G. Schaur (2013). "Time as a Trade Barrier," *American Economic Review* 103: 2935-2959.

Rosenbaum, P. and D. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70: 41-55.

Volpe Martincus, C. and J. Carballo (2008). "Is Export Promotion Effective in Developing Countries? Firm-level Evidence on the Intensive and Extensive Margins of Export Growth," *Journal of International Economics* 76: 89-106.

Volpe Martincus, C., Carballo, J., and A. Graziano (2015). "Customs," *Journal of International Economics* 96: 119-137.

WTO (World Trade Organization). 2013. *Fourth Global Review of Aid for Trade 2013: Connecting to Value Chains.* Geneva: World Trade Organization.

Appendix A. Conditioning Variables Definition

Table A1. Firm Covariates Whose- Pre-Treatment Values are Used in the Synthetic Control Method Estimation

EU share	share of firm's total value of imports in a month accounted for by imports from the EU 27 countries.
Special customs clearance	share of firm's total value of imports in a month accounted for by declarations that benefit from entering Serbia under a special customs clearance regime other than the in-house clearance.
Share of imports of:	**share of firm's total value of imports in a month accounted for by:**
Animal and vegetable products and foodstuffs	HS 10-digit products belonging to HS2 chapters 01-24
Minerals	HS 10-digit products belonging to HS2 chapters 25-27
Chemicals, plastics and rubbers	HS 10-digit products belonging to HS2 chapters 28-40
Raw hides, skins, leathers, furs, textiles and footwear	HS 10-digit products belonging to HS2 chapters 41-43 and 50-67
Wood and paper products	HS 10-digit products belonging to HS2 chapters 44-49
Stone and glass	HS 10-digit products belonging to HS2 chapters 68-71
Metals	HS 10-digit products belonging to HS2 chapters 72-83
Machines and electrical	HS 10-digit products belonging to HS2 chapters 84-85
Transportation	HS 10-digit products belonging to HS2 chapters 86-89
Miscellaneous and services	HS 10-digit products belonging to HS2 chapters 90-99

Appendix B. Synthetic Control Method Detailed Results Tables

Table B1. SCM Detailed Results for All IHC Firms - Clearance Time at 75th percentile

IHC firm #	6-month average treatment effect: change in log hours at 75th percentile of firm-month distribution ($\beta\_\widehat{hrs}\_75_f$)	Number of donor firms used to construct synthetic control firm	Rank of IHC firm relative to 25 placebo firms	Percentile rank statistic (rank/26)
1	0.066	11	17	0.654
2	-0.814	11	3	0.115
3	-0.183	8	6	0.231
4	-0.401	11	7	0.269
5	-0.465	46	4	0.154
6	0.326	5	18	0.692
7	-0.545	10	4	0.154
8	-1.386	9	2	0.077
9	-0.853	7	2	0.077
10	-0.279	10	9	0.346
11	0.274	22	24	0.923
12	-1.013	8	3	0.115
13	-0.577	10	3	0.115
14	-0.780	13	2	0.077
15	0.593	42	23	0.885
16	-0.137	8	13	0.500
17	-0.028	6	11	0.423
18	-0.398	4	7	0.269
19	-1.360	6	1	0.038
20	-0.247	7	8	0.307
21	-0.896	9	2	0.077
Mean % change in 75th pct. clearance time	-0.434		Mean percentile rank statistic ($\bar{p}$)	0.310
Median % change in 75th pct. clearance time	-0.401			

Notes: Synthetic control firms were constructed with the benchmark specification that includes three pre-treatment values of the outcome variable as part of the set of conditioning variables. The summary statistics at the bottom left of the table are obtained as the mean and the median of the treatment effect across the 21 firms. The statistic at the bottom right of the table $\bar{p}$ given 21 firms has 5% critical values of 0.377 and 0.623 and 1% critical values of 0.339 and 0.660 (based on Appendix table A1 in Dube and Zipperer (2015)).

Table B2. SCM Detailed Results for All IHC Firms - Clearance Time at 90th percentile

IHC firm #	6-month average treatment effect:change in log of clearance time at 90th percentile of firm-month distribution ($\widehat{\beta\_hrs\_90}_f$)	Number of donor firms used to construct synthetic firm	Rank of IHC firm relative to 25 placebo firms	Percentile rank statistic (rank/26)
1	0.189	10	18	0.692
2	-1.613	10	1	0.038
3	-0.125	9	11	0.423
4	-0.362	13	10	0.385
5	-0.122	9	12	0.462
6	0.262	6	18	0.692
7	-0.580	8	10	0.385
8	-2.088	6	1	0.038
9	-0.377	9	5	0.192
10	-0.048	5	12	0.462
11	0.918	31	25	0.962
12	-1.107	8	4	0.154
13	-0.845	8	3	0.115
14	-0.682	20	4	0.154
15	0.513	50	21	0.808
16	0.023	6	15	0.577
17	0.334	9	18	0.692
18	-0.802	9	2	0.077
19	-0.762	5	4	0.154
20	-0.290	6	8	0.308
21	-1.041	8	4	0.154
Mean % change in 90th pct clearance time	-0.410		Mean percentile rank statistic ($\bar{p}$)	0.377
Median % change in 90th pct clearance time	-0.362			

Notes: Synthetic control firms were constructed with the benchmark specification that includes three pre-treatment values of the outcome variable as part of the set of conditioning variables. The summary statistics at the bottom left of the table are obtained as the mean and the median of the treatment effect across the 21 firms. The statistic at the bottom right of the table $\bar{p}$ given 21 firms has 5% critical values of 0.377 and 0.623 and 1% critical values of 0.339 and 0.660 (based on Appendix table A1 in Dube and Zipperer (2015)).

Table B3. SCM Detailed Results for All IHC Firms – Inspection rates

IHC firm #	6-month average treatment effect: change in inspection rate ($\widehat{\beta\_inspect}_f$)	Number of donor firms chosen by synthetic control procedure	Rank of IHC firm relative to 25 placebo firms	Percentile rank statistic (rank/26)
1	0.037	10	22	0.846
2	-0.058	14	6	0.231
3	0.000	10	11	0.423
4	0.002	11	17	0.653
5	-0.008	14	13	0.500
6	0.009	8	20	0.769
7	0.001	12	14	0.538
8	-0.079	57	3	0.115
9	-0.023	14	7	0.269
10	-0.008	50	15	0.577
11	-0.014	12	11	0.423
12	0.008	10	16	0.615
13	-0.021	11	7	0.269
14	-0.026	9	8	0.307
15	-0.022	6	9	0.346
16	0.023	10	18	0.692
17	-0.014	9	9	0.346
18	0.033	5	21	0.808
19	-0.005	7	14	0.538
20	0.001	4	12	0.462
21	-0.007	7	11	0.423
Mean % change in inspection rate	-0.008		Mean percentile rank statistic ($\bar{p}$)	0.483
Median % change in inspection rate	-0.007			

Notes: Synthetic control firms were constructed with the benchmark specification that includes three pre-treatment values of the outcome variable as part of the set of conditioning variables. The summary statistics at the bottom left of the table are obtained as the mean and the median of the treatment effect across the 21 firms. The statistic at the bottom right of the table $\bar{p}$ given 21 firms has 5% critical values of 0.377 and 0.623 (based on Appendix table A1 in Dube and Zipperer (2015)).

Table B4. SCM Detailed Results for All IHC Firms – Import Value

IHC firm #	6-month average treatment effect: change in log import value ($\beta\_\widehat{import}_f$)	Number of donor firms chosen by synthetic control procedure	Rank of IHC firm relative to 25 placebo firms	Percentile rank statistic (rank/26)
1	0.404	9	24	0.923
2	0.121	9	21	0.808
3	-0.002	7	14	0.538
4	0.048	9	15	0.577
5	0.236	7	22	0.846
6	0.104	8	17	0.654
7	0.387	68	21	0.808
8	-0.448	9	2	0.077
9	0.370	9	21	0.808
10	-0.197	12	5	0.192
11	0.323	8	20	0.769
12	-0.163	8	6	0.231
13	-0.652	8	1	0.038
14	-0.126	9	10	0.385
15	0.147	7	20	0.769
16	0.365	4	24	0.923
17	0.106	6	20	0.769
18	-0.224	2	9	0.346
19	-0.778	10	2	0.077
20	0.520	9	24	0.923
21	0.368	11	23	0.885
Mean % change in import value	0.043		Mean percentile rank statistic ($\bar{p}$)	0.588
Median % change in import value	0.106			

Notes: Synthetic control firms were constructed with the benchmark specification that includes three pre-treatment values of the outcome variable as part of the set of conditioning variables. The summary statistics at the bottom left of the table are obtained as the mean and the median of the treatment effect across the 21 firms. The statistic at the bottom right of the table $\bar{p}$ given 21 firms has 5% critical values of 0.377 and 0.623 (based on Appendix table A1 in Dube and Zipperer (2015)).

Appendix C. Synthetic Control Method: Illustration for a Single IHC Firm

As an illustration of the SCM estimator we show the analysis of the effects of the IHC program with a graphical representation for one of the IHC-adopting firms. Our purpose in this appendix is pedagogical, so we choose a firm for which the SCM produces visual outcomes that are instructive across all outcome variables. Graphs for all treated firms can be made available to interested readers. We first illustrate simple SCM results for log median clearance time. Using this outcome variable and this IHC example firm, we move to an illustration of placebo-based inference for a single firm. Using the same IHC example firm and outcome variable, we illustrate our alternative specifications to constructing the synthetic control firm. Finally we show, for the same IHC example firm, SCM results for our other outcome variables, log clearance times at the 75th and 90th percentile, inspection rates, and log import value. These latter estimates use our benchmark specification for fitting the SCM procedure.

Recall that for the estimation of the effect of the IHC program on time to clear customs, the SCM procedure chooses vectors of weights V^* and W^* to jointly minimize a) the squared difference of treated firm and synthetic control firm in the outcome variable in the pre-treatment period, and b) the difference between the average conditioning characteristics of the treated firm and the synthetic control firm. In the post-treatment months that follow the firm's first use of the IHC program, firm weights W^* are fixed and the procedure tracks the outcomes for each IHC firm and its synthetic control counterpart as we illustrate below.

Figure C1 shows the evolution of the moving average median time to clear customs for the chosen IHC example firm in the solid line and for its synthetic control firm in the dashed line. The vertical dotted line indicates the month in which this firm began to use the IHC program. In the pre-treatment period, log median clearance times range from -0.5 to 0.5 (in levels this implies the firm's monthly clearance times range from about 0.5 to 2.5 hours). The fitting procedure used to produce our benchmark estimates includes among the conditioning variables the outcome variable's values in periods 1, 10 and 18 of the pre-treatment period. The synthetic control firm is constructed to follow the dynamics of the treated firm over the pre-treatment period as well as a bundle of import characteristics, but the procedure puts a premium on hitting the outcome of the treated firm for those three specific points in time.

A striking fact about Figure C1 is that the clearance time for the treated firm falls dramatically after the IHC program is adopted and is consistently stable thereafter. The stability of the clearance time at the post-treatment level indicates that the treated firm is now clearing at least half of its shipments under the IHC program in every month; the default clearance time of 0.5 hours is observed as the median in every month following treatment. While this example firm clears more than half its shipments at the default value post treatment, this is not true of all the firms that adopt the IHC program.

Estimates of the treatment effect of the IHC program are shown in the period to the right of the dotted line. The gap between the line representing outcomes for the synthetic control firm and the line representing outcomes for the treated firm represents the estimated time savings in each month that arise because the firm adopted the IHC program. In our formal notation, the gap at any point in time is $\hat{\alpha}_{1t}$. Because median clearance times for the treated firm fall dramatically while staying relatively stable for the synthetic control firm, the monthly estimated treatment effects for this IHC example firm are negative and at values indicating approximately 1 log point. This estimate does not take into account statistical uncertainty, and it represents the outcome for a

single firm. Outcomes are more ambiguous for other firms, and estimated treatment effects can vary according to the method used to construct the synthetic control firm.

Figure C1. Median Clearance Time for IHC Example Firm and its Synthetic Control Firm
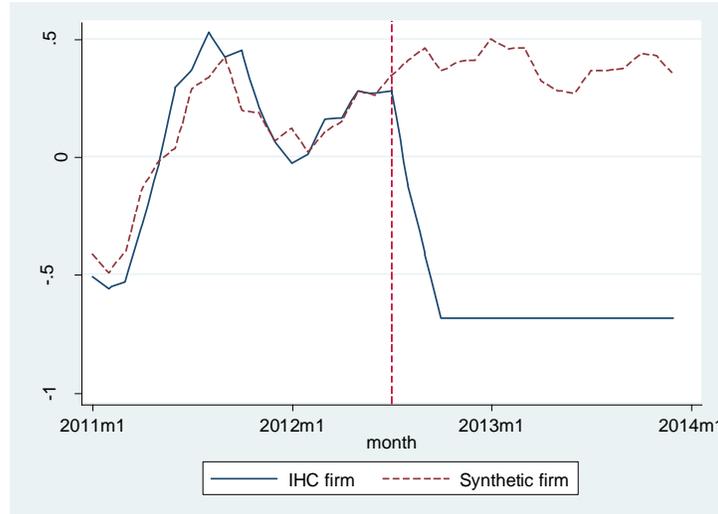
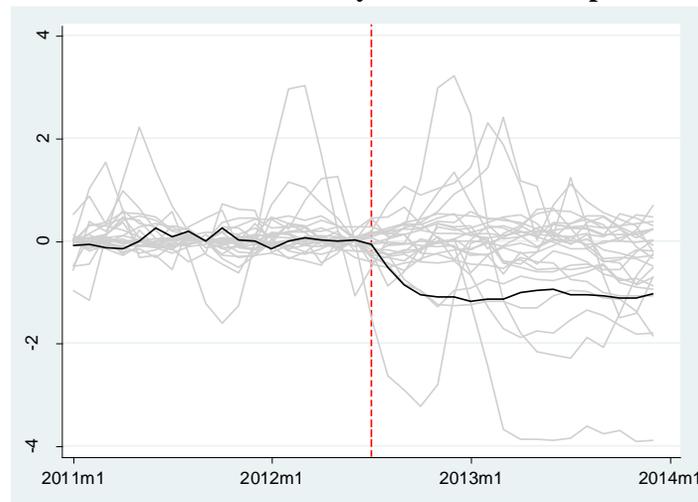
Source: Authors' estimates based on Serbian customs data.
Notes: the outcome variable shown in the Y-axis is the 3-month moving average of log median hours for customs clearance for the treated firm and the synthetic control firm. The vertical line indicates the date when the treated firm began to use the IHC program.

Formal analysis of whether or not the gap between the treated firm and the synthetic firm in Figure C1 represents a real reduction in median customs clearance time for the treated firm requires a placebo test. While the gap for this particular IHC example firm is quite striking, for other firms the treatment effect is not always negative, nor is its sign stable over time. Another lesson from Figure C1 is that the scale of the Y-axis is quite small. The IHC example firm clears most of its imports in less than 2.5 hours even before adopting the program, and the SCM procedure constructs a synthetic control firm with similar low clearance times. This underscores the important point that the firms adopting the IHC program are not typical, and usually clear customs for their imports quickly even before adopting the program.

Using the notation from Section 3.2, we designate the estimated treatment effect in each post-treatment period t for IHC example firm f as $\hat{\alpha}_{ft}$, which measures the visible "gap" between the IHC example firm and its synthetic control firm. Figure C2 shows the gap that is consistent with the observed outcome in Figure C1: it fluctuates around zero for the pre-treatment period, but rises in the post-treatment period.

The gap opening up between the IHC example firm and the synthetic control firm may be attributable to the IHC program, but it may also represent random shocks that separate the IHC example firm and the synthetic control firm post-treatment. In order to put the observed gaps in context, we conduct a placebo analysis drawing 25 firms that do not qualify for IHC clearance at random (from within the same decile of import value as the IHC example firm). We apply the SCM to each placebo firm, constructing a synthetic control firm that matches the placebo firm in terms of the dynamics of the outcome variable and the conditioning firm characteristics. We apply the same treatment date for the 25 placebo firms as for the corresponding IHC example firm, even though the placebo firms never participate in the IHC program. This procedure generates a distribution of random outcomes for the treatment effect against which the observed values of $\hat{\alpha}_{ft}$

can be compared. An example of this placebo analysis is provided in Figure C3, where the darker gap line for the IHC example firm (the same as shown in Figure C2 but with a Y-axis with a different scale to be able to show the variation from the placebo firms) is plotted along with the lighter gap lines for its 25 placebo firms. For much of the post-treatment period the gap line for the IHC example firm lies in the lower part of the distribution of placebo outcomes, but it is not clearly outside it. This means that in general we do not have enough evidence to suggest that the negative estimated treatment effect for this particular firm is attributable to anything other than random variation.

Figure C2. Gap in Median Clearance Times for IHC Example Firm
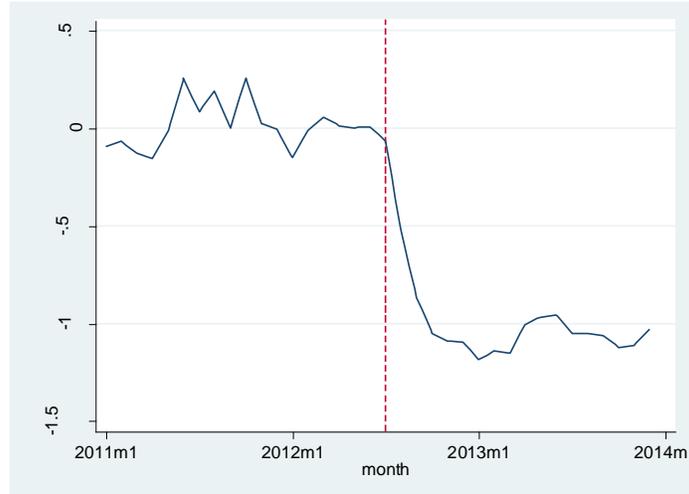
Source: Authors' estimates based on Serbian customs data.
Notes: the outcome variable shown in the Y-axis is the gap between the 3-month moving average of log median hours of clearance time for the treated firm and that for the synthetic control firm. The vertical line indicates the date when the treated firm began to use the IHC program.

But Figure C3 also suggests a possibility that we explore in our main results in the text,- that an analysis based on the treatment effect for a single treated firm simply lacks the statistical power necessary to conclude that the IHC program had an impact. Figure C3 shows that the gap for the IHC example firm consistently lies in the lower half of the distribution of the gaps for placebo firms. If this pattern is sufficiently consistent across the firms that adopted the IHC program, then one could conclude the IHC program had a negative effect on firm median customs clearance time. Dube and Zipperer (2015) apply such logic in their study of minimum wage increases in U.S. states.

The SCM estimates summary in Table 10 shows that for median clearance time there is indeed a tendency for the summary statistics to lie in the bottom half of the distribution of placebos. The median and Hodges-Lehmann point estimates both suggest that a negative impact of the IHC program on clearance time is more likely than a positive one. But the formal tests for statistically significant effects still fail for median clearance time. The variability of the placebo estimates is sufficiently large that the tendency for estimated treatment effects of IHC to lie in the bottom half of the distribution does not generate a statistically significant result.

Figure C3. Median Clearance Time Placebo Analysis for IHC Example Firm
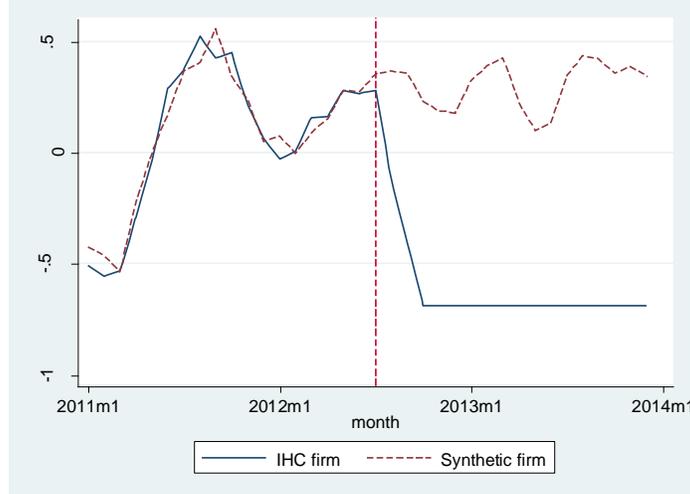

Source: Authors' estimates based on Serbian customs data.
Notes: the variable shown in the Y-axis is the gap between the 3-month moving average of log median clearance time for the treated firm and that for the synthetic control firm. The darker line shows the gap for the treated firm while the lighter lines show the gap for each of the 25 placebo firms. The vertical line indicates the date when the treated firm began to use the IHC program.

Next we turn to an illustration of the two alternative approaches used to construct the synthetic control firm – discussed in the robustness checks in Section 6 - using the same IHC example firm and log median clearance time as the outcome variable.

The first approach includes all pre-treatment lagged values of the outcome variable in the set of conditioning variables. Including these variables requires the SCM estimator to put a high priority on replicating the dynamics of the outcome variable in the pre-treatment period. In practice, this specification discounts entirely the other conditioning characteristics (such as the EU share or the share of specific sectors in pre-treatment imports); the v-weights on these variables go to zero. In effect, these other characteristics play no role whatsoever in fitting when all the pre-treatment lagged values of the outcome variable are included as conditioning variables. The results from constructing a synthetic control firm using this approach are illustrated in Figure C4. This exercise suggests a negative treatment effect from the IHC program on median clearance times for this IHC example firm. However, the placebo analysis (not shown here) indicates sufficient variability amongst estimated treatment effects for the placebo firms to suggest a lack of statistical significance for the negative treatment effect that is visible here.

Figure C4. Median Clearance Time for IHC Example Firm and its Synthetic Control Firm - Using All Pre-Treatment Values of Median Clearance Time

The second approach includes only the average pre-treatment value of the outcome variable as part of the conditioning variables and thus removes any 'help' given to the SCM procedure in hitting dynamics in the outcome variable before the treatment. The average pre-treatment outcome pins down the approximate scale of the clearance time variable, but only the characteristics of the firm's import bundle are allowed to inform dynamic movements. As Figure C5 indicates, when this approach is used to construct a synthetic control firm there is a relatively poor fit of the dynamics in the pre-treatment period. It does not appear that the dynamics of median clearance time are captured well by the (average pre-treatment) composition of the firm's import bundle. But the estimates of the treatment effect in this case are similar to those obtained with the benchmark method of constructing the synthetic control firm. Once again, this robustness check suggests a negative estimated treatment effect of the IHC program on the IHC example firm's monthly median clearance time.

Next we return to the benchmark approach for constructing the synthetic control firm, but present the results for two other outcome variables, the log of the 75th percentile and the log of the 90th percentile of the distribution of customs clearance times. Figure C6 shows that the pre-treatment values are somewhat higher for the 75th percentile of clearance times than for the median (shown in Figure C1). After treatment, the 75th percentile of clearance times for this IHC example firm is very near the IHC default clearance time of 0.5 hours, and sometimes at it. Because the values of the 75th percentile of clearance times for the synthetic control firm are higher than in the case of the median this implies larger values of $\hat{\alpha}_{ft}$. This fact for the IHC example firm is reflected in the larger observed treatment effects at the 75th percentile for all firms described in Section 5.
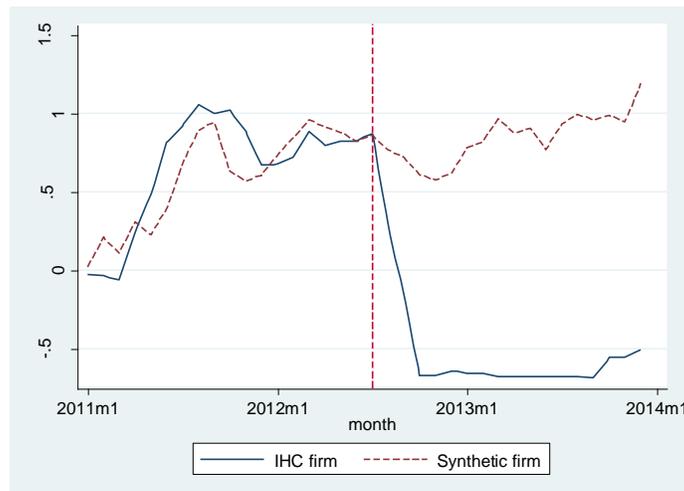
Figure C5. Median Clearance Time for IHC Example Firm and its Synthetic Control Firm – Using Average Pre-Treatment Value of Median Clearance Time
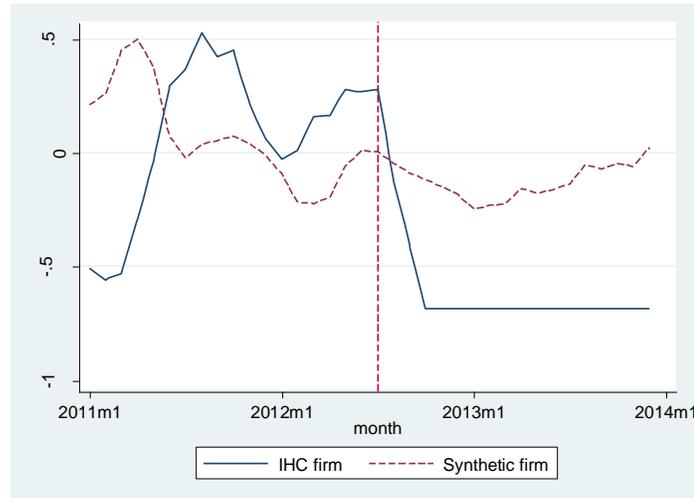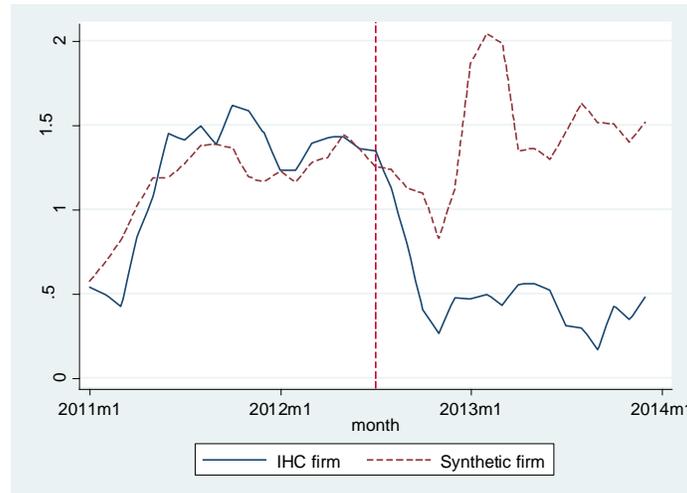

Source: Authors' estimates based on Serbian customs data.
Notes: the outcome variable shown in the Y-axis is the 3-month moving average of log median hours for customs clearance for the treated firm and the synthetic control firm. The vertical line indicates the date when the treated firm began to use the IHC program.

Figure C6. Clearance Time at the 75th Percentile for IHC Example Firm and its Synthetic Control Firm
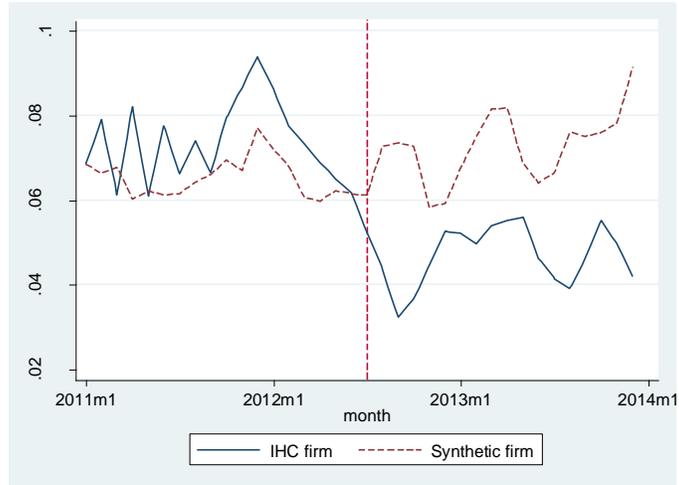

Source: Authors' estimates based on Serbian customs data.
Notes: the variable shown in the Y-axis is the gap between the 3-month moving average of log of the 75th percentile of clearance times for the treated firm and the synthetic control firm. The vertical bar indicates the date when the treated firm adopted the IHC program.

Figure C7 shows that for this IHC example firm the 90th percentile of clearance times rarely takes the default value. Pre-treatment clearance times are, of course, higher at the 90th percentile than at the 75th or at the median. But because the IHC example firm does not clear 90 percent of its shipments through IHC, its post-treatment values of the 90th percentile of clearance times do not fall to the IHC default value. The values for the synthetic control firm are also more volatile. These results are useful for understanding why the size of the estimated treatment effect is smaller at the 90th percentile than at the 75thpercentile, and why the variability of the estimated effect is larger.

Figure C7. Clearance Time at the 90th Percentile for IHC Example Firm and its Synthetic Control Firm

Figure C8 shows a graphical representation of the benchmark estimates for the same IHC example firm with the inspection rate as the outcome variable. The inspection rate is somewhat volatile (although this volatility is observed on a very small scale). The synthetic control firm matches the approximate scale of the IHC example firm but not its volatility. Following treatment the inspection rate falls for this IHC example firm, in absolute terms and relative to the synthetic control firm. These effects are not significant in the placebo analysis. Neither are the effects significant when we pool across IHC firms, as seen in Section 5.

Finally, Figure C9 shows a graphical representation of the benchmark estimates for the same IHC example firm and log import value as outcome variable. Before treatment, log import value for the synthetic control firm matches the approximate scale and dynamics of the IHC example firm. The estimates for this IHC example firm suggest a positive treatment effect of the IHC program on log import value, but this effect is not significant according to the unreported placebo analysis. Neither are the effects significant when we pool across IHC firms, as discussed in Section 5.
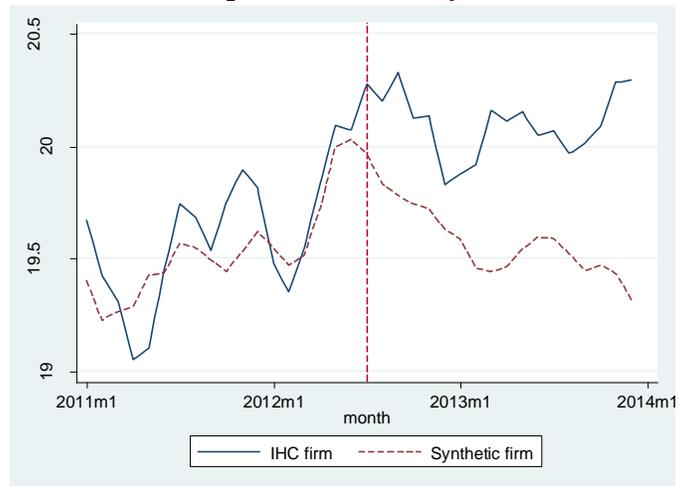
Figure C8. Inspection Rate for IHC Example Firm and its Synthetic Control Firm

Source: Authors' estimates based on Serbian customs data.
Notes: the outcome variable shown in the Y-axis is the 3-month moving average of the inspection rate for the treated firm and the synthetic control firm. The vertical line indicates the date when the treated firm began to use the IHC program.

Figure C9. Import Value for IHC Example Firm and its Synthetic Control Firm

Source: Authors' estimates based on Serbian customs data.
Notes: the outcome variable shown in the Y-axis is the 3-month moving average of log import value for the treated firm and the synthetic control firm. The vertical line indicates the date when the treated firm began to use the IHC program.