Institute for Fiscal Studies

Vladimir Ponczek
Gabriel Ulyssea

Working paper 21/08

Enforcement of labor regulation and the labor market effects of trade: evidence from Brazil

Economic and Social Research Council

Enforcement of Labor Regulation and the Labor Market Effects of Trade: Evidence from Brazil*

Vladimir Ponczek
EESP/FGV

Gabriel Ulyssea†
UCL, IFS, CEPR and IZA

March 22, 2021

Abstract

How does enforcement of labor regulations shape the labor market effects of trade? Does the informal sector introduce greater *de facto* flexibility, reducing employment losses during bad times? To tackle these questions, we exploit local economic shocks generated by trade liberalization and variation in enforcement capacity across local labor markets in Brazil. In the aftermath of the trade opening, regions with stricter enforcement observed: (i) lower informality effects; (ii) larger losses in overall employment; and (iii) greater reductions in the number of formal plants. Regions with weaker enforcement observed opposite effects. All these effects are concentrated on low-skill workers. Our results indicate that greater *de facto* labor market flexibility introduced by informality allows both formal firms and low-skill workers to cope better with adverse labor market shocks.

*This research is supported by the R4D initiative on Employment and labour market outcomes, funded by the Swiss National Science Foundation and the Swiss Development Cooperation. We are thankful to Penny Goldberg, Rafael Dix-Carneiro, Brian Kovak, and Brian McCaig, as well as participants at various conferences and seminars for helpful comments and discussions. All remaining errors are ours.

†Contact author. Address: Department of Economics, University College London. Drayton House, 30 Gower Street, London WC1H 0AX, UK. Email: g.ulyssea@ucl.ac.uk

1 Introduction

Many developing countries, most notably in Latin America, underwent major trade liberalization episodes in the 1980s and early 1990s (Goldberg and Pavcnik, 2007). Despite the many expected gains from trade, concerns about negative labor market consequences have always been present in these countries.[1] In particular, one major concern is that trade opening could induce a reallocation from formal to informal jobs, specially among less skilled workers (Goldberg and Pavcnik, 2003). Since informal jobs are typically of lower quality and are not covered by labor regulations nor social security, this informality effect could represent a large welfare loss from trade opening.

However, informality also introduces greater *de facto* labor market flexibility, which can be particularly relevant in the presence of burdensome and strict labor regulations. Higher flexibility may help firms and workers to cope better with negative economic shocks, which could reduce employment losses relative to a counterfactual scenario with perfect enforcement and no informality. This conjecture has important implications for how one interprets the labor market effects from trade and their potential consequences for welfare. More broadly, it implies that the rigidity introduced by labor market regulations can lead to worse labor market outcomes and potentially amplify employment losses from adverse economic shocks. This latter point directly speaks to the extensive literature that analyzes the consequences of labor regulations and labor market rigidity for labor market performance.[2]

This paper tackles these issues by exploiting Brazil's large scale, unilateral trade liberalization episode of the early 1990's. Brazil is an attractive empirical setting for at least three reasons. First, the unilateral trade liberalization had substantial and heterogeneous effects across local labor markets (e.g. Kovak, 2013; Dix-Carneiro and Kovak, 2017, 2019). Second, just before the beginning of the trade liberalization process, in 1988, Brazil underwent a major Constitutional reform that substantially increased the restrictiveness and the direct costs associated to labor regulation (Barros and Corseuil, 2004).[3] Third, enforcement of labor regulation varies greatly across regions in Brazil (Almeida and Carneiro, 2012). We exploit the geographic variation in the intensity of trade shocks and enforcement of labor regulation to assess if, and to what extent, the presence of stricter enforcement of a costly regulatory framework shapes the labor market

[1] A growing literature consistently documents that local economies that become more exposed to foreign competition observe worse labor market outcomes relative to those that are less exposed (e.g. Autor et al., 2013; Kovak, 2013; Costa et al., 2016; Dix-Carneiro and Kovak, 2017, 2019).

[2] This is an extensive literature, which we discuss ahead.

[3] According to the employment index in Botero et al. (2004), the cost of labor regulation in Brazil is 20 percent above the mean and median of 85 countries in the world and more than 2.5 times as large as in the United States.

responses to trade liberalization. More broadly, we investigate whether greater *de facto* labor market flexibility (introduced by informality) leads to lower employment losses in face of an adverse economic shock.

To get to these questions, we construct a measure of local, trade-induced shocks based on changes in tariffs at the industry level combined with the initial sectoral composition of employment across regions, which remains fixed at the levels observed before the trade opening process started (e.g. Topalova, 2010; Kovak, 2013). An important identification assumption is the (conditional) exogeneity of these trade shocks relative to (unobserved) pre-existing trends in local labor markets. Previous papers that use the same regional trade shock document direct evidence in support of this assumption (e.g. Kovak, 2013; Dix-Carneiro and Kovak, 2017, 2019). A second key aspect of our empirical strategy is the measurement of enforcement capacity and intensity across local economies. In Brazil, enforcement of labor regulation is the sole responsibility of the Ministry of Labor and the technology of enforcement is quite straightforward: labor inspectors are assigned to labor offices (L.O.) located in municipalities across the country and they travel by car to inspect firms (Almeida and Carneiro, 2012). Hence, greater distances to L.O. imply that firms are less likely to be inspected and enforcement is more likely to be weak (all things equal). We thus use distance to the nearest labor office as a proxy for enforcement capacity in a given local market. We collect new data on the date of creation of all labor offices in Brazil and restrict the analysis to those created before the trade opening process started. Hence, our measure of enforcement capacity is pre-determined with respect to future trade shocks and local labor market conditions.

We start by examining the basic effects of regional trade shocks on labor market outcomes. We use individual-level Census data from 1991 and 2000 to compute local labor market outcomes net of the influence of socio-demographic variables (e.g. gender, schooling and age). Our results confirm the findings of previous studies: between 1991 and 2000, regions more exposed to the trade liberalization shock experienced a substantial increase in both informality and non-employment relative to regions less exposed.[4] The novelty here comes from the analysis of these effects across skill levels, which shows striking results. Almost all negative effects on informality, non-employment and wages come from low-skill workers. High-skill workers show no informality effects and much smaller and marginally significant effects on non-employment and wages.

We then move to the focus of the paper, which is the analysis of heterogeneous effects

[4] Dix-Carneiro and Kovak (2019) also show that in the longer run (until 2010) the effects on non-employment vanish but on informality persist (and even amplify). We focus on the period up to 2000 because our focus lies on investigating the extent to which stricter enforcement reduces/amplifies employment effects in the aftermath of a negative demand shock.

across enforcement capacity levels, which is proxied by the maximum driving distance to the nearest labor office. The results show that regions with higher enforcement capacity observe lower informality but greater non-employment effects as a response to the trade shock. Symmetrically, regions with lower enforcement capacity experience greater informality but lower non-employment effects. Again, all the effects are concentrated on low-skill workers, with no heterogeneous effects among high-skill workers. The magnitudes of these effects are large. For a strong local trade shock (tariff reduction of 0.1 log points), a region with very low enforcement capacity (90th percentile of the distance distribution) would experience an increase of 10 p.p. on informality, but nearly zero effects on non-employment. In contrast, a region with high enforcement capacity (10th percentile of the distance distribution) would experience a 3 p.p. increase in informality but a much stronger increase in non-employment, of 3.9 percentage points. As for wages, the effects are large in magnitude and go in the expected direction, i.e. when enforcement is weaker there are greater wage losses. However, the coefficients are imprecisely estimated and we find no statistically significant heterogeneous effects on wages for neither skill group in our benchmark specifications.

To complement these results from the Demographic Census, we use administrative data from the Ministry of Labor that contain the universe of formal firms and workers. We find that regions with weaker enforcement experience a stronger reduction in formal employment relative to regions with greater enforcement capacity. This indicates that, beyond preserving jobs, there is also a "switching effect" from formal to informal jobs in harder-hit regions with weaker enforcement. However, we show that the higher flexibility introduced by weaker enforcement also leads to greater survival of formal establishments, as regions with weaker enforcement capacity have smaller losses in the number of formal plants. This result is consistent with the fact that a large fraction of informal employment is located in formal firms, the so-called intensive margin of informality (Ulyssea, 2018). Moreover, it indicates that the intensive margin plays an important role in formal firms' survival in face of an adverse economic shock.

In sum, our results show that, in the years following the unilateral trade liberalization, regions with stricter enforcement (and therefore higher labor market rigidity) experienced the following: (i) less switching from formal to informal jobs and lower overall informality effects; (ii) larger employment losses; and (iii) greater reductions in the number of formal plants. The opposite is observed in regions where enforcement capacity is weak and *de facto* labor market flexibility is high: there are strong informality effects, but no statistically significant effects on employment and greater survival of formal firms. These effects are completely driven by low-skill workers. Our results therefore indicate that greater *de facto* labor market flexibility introduced by informality allows both formal firms and low-skill workers to cope better with adverse labor market shocks. Put differently, the results suggest that informality acts as an employment buffer in face of negative economic shocks, but this seems to be the case only for low-skill workers.

Even though our measure of enforcement capacity is pre-determined, one potential concern about the empirical strategy is that the location of labor offices is obviously not random. Thus, the enforcement capacity measure could be capturing the effects of other characteristics that are not accounted for in our specifications. For example, distance to labor offices could be capturing how remote a given area is, i.e. less connected to important economic centers. More remote regions may also have a higher proportion of low-productivity firms, which are more likely to respond to negative shocks by increasing informal employment. We address this and other potential threats to identification in different ways. We start by noting that all regressions are estimated in first difference, therefore accounting for time-invariant, unobserved local economies' characteristics. Our benchmark specifications also allow for state-specific trends and for differential trends across micro-regions with different initial demographic conditions within states.

Hence, the threat to identification would have to come from some omitted, time-varying, regional characteristic that drives labor market outcomes within a given state, which is not captured by local economies' initial socio-demographic conditions, and which is also correlated with the location of labor offices. The extensive robustness analysis in Section 4.3 shows that this is unlikely to be the case, as our results are robust to a variety of additional potential confounders. We start by investigating whether reversion to the mean across regions with lower and higher initial levels of informality and non-employment could be driving our results. This would be the case if our enforcement capacity measure was capturing heterogeneity in initial informality levels, rather than enforcement capacity *per se*. This is not the case, as the results are robust to the inclusion of 1980's informality and non-employment rates. We also investigate whether our results are capturing the "remoteness effect" mentioned above. We do so by including the median driving distance to the state's capital, thus allowing for differential trends by proximity to the capital, and the results remain unchanged (if anything become stronger). Finally, we control for local government per capita spending and the Gini coefficient, both measured in 1991. The former aims to control for the possibility that distance to the labor offices could be a proxy for availability of local public goods and infra-structure in general, rather than enforcement capacity *per se*. The latter controls for the initial level of inequality in the region, which is also a relevant indicator of local economic conditions. Our results remain unchanged in both cases.

In the final part of the paper, we turn our focus to a more direct measure of enforcement intensity, which is the total number of inspections per formal firms in a given local market. This measure is potentially subject to measurement error and it is likely to be an endogenous regressor. We thus use distance to the nearest L.O. as an instrument for enforcement in a limited information maximum likelihood estimator. Again considering a high intensity local trade shock (tariff reduction of 0.1 log points), low-skill workers in a region with weak enforcement – 0.9 inspections per 100 firms (the 10th percentile) – would experience an increase of 12.1 percentage points in informality but no disemployment effects. In contrast, low-skill workers in a region with strict enforcement – 17.2 inspections per 100 firms (the 90th percentile) – would experience no informality effects but an increase of 10.3 percentage points in non-employment rates. It is worth noting that the regional trade shock is associated to an increase of 5.2 and 2.7 percentage points in informality and non-employment among low-skill workers, respectively. Hence, the strength of enforcement (or lack thereof) can lead to labor market responses in both informality or non-employment that are substantially larger than the average effect from the trade shock.

Our paper contributes to three literature streams. First, the literature on trade and local labor markets, which includes (but is not restricted to) Topalova (2010), Kovak (2013), Autor et al. (2013), Hakobyan and McLaren (2016), and Dix-Carneiro and Kovak (2017, 2019). Second, the literature that analizes the relationship between trade opening and informality (Goldberg and Pavcnik, 2003; Menezes-Filho and Muendler, 2011; Bosch et al., 2012; Dix-Carneiro and Kovak, 2017, 2019). In contrast with both literature streams, we focus on a new dimension: the interaction between the enforcement of labor regulation and trade policies, and how these interactions shape the labor market adjustment to trade shocks. Also importantly, we document that the effects of trade opening on informality and non-employment are mostly concentrated on low skill workers, both on average and across different enforcement levels. Finally, our paper also dialogues with the literature that argues that informality introduces *de facto* flexibility to otherwise very rigid formal labor markets that are subject to burdensome and costly regulatory frameworks (e.g. Meghir et al., 2015; Ulyssea, 2018).[5] More broadly, our results speak to the extensive literature about the consequences of labor regulations and labor market rigidity (e.g. Nickell and Layard, 1999; Botero et al., 2004; Nickell et al., 2004; Besley and Burgess, 2004; Tella and MacCulloch, 2005; Freeman, 2010; Adhvaryu et al., 2013; Almeida and Poole, 2017).

The remainder of the paper is structured as follows. Section 2 briefly describes the

[5]In an earlier paper, Boeri and Garibaldi (2005) argue that despite advances in monitoring capacity by the government, informality is "tolerated" because it attenuates unemployment. Similarly, Ulyssea (2010) quantitatively shows in a two-sector matching model that higher enforcement leads to higher unemployment and lower welfare. Even though our empirical results refer to a very different setting, they are consistent with the mechanisms highlighted in both papers.

trade liberalization process and the measure of local trade shock used, the institutional background and data. Section 3 describes the empirical strategy, while Section 4 presents the empirical results. Section 5 concludes.

2 Background and Data

2.1 Trade Liberalization and Local Trade Shocks in Brazil

Until 1990, Brazil was characterized by a complex system of protection against foreign competition that included both tariff and non-tariff barriers (Kume et al., 2003). In 1988-1989, there was a first move towards reforming the structure of protection, which reduced tariff redundancy and special regimes, among other measures. In March 1990, the newly elected president unexpectedly eliminated non-tariff barriers, typically replacing them with higher import tariffs in a process known as "tariffication". This implied that, from 1990 onwards, tariffs started to accurately reflect the actual level of protection faced by Brazilian industries.[6]

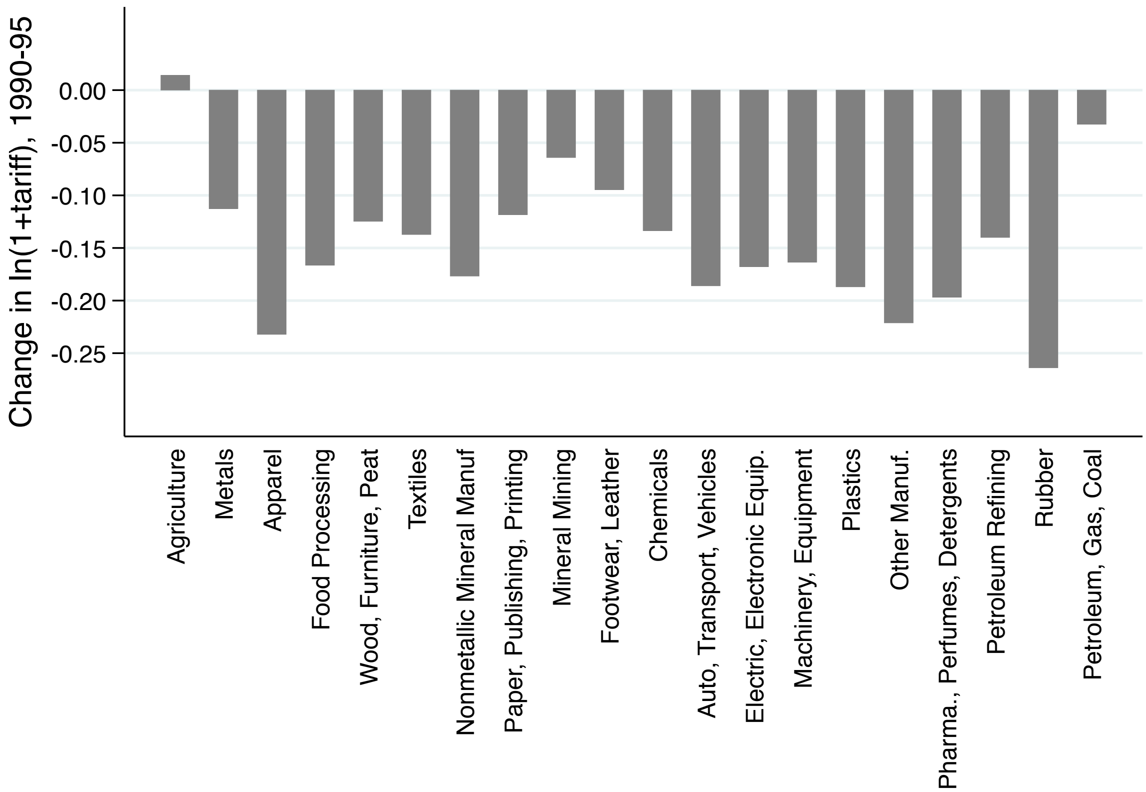
From 1990 until 1995, Brazil implemented a major unilateral reduction in trade tariffs. During this period, the average tariff fell from 30.5 percent to 12.8 percent and the standard deviation fell from 14.9 percent to 7.4 percent.[7] Hence, not only the overall level of protection decreased but also the variation across industries was substantially reduced. Figure 1 shows the percentage change in tariffs across the main industries, which is one of the sources of variation we exploit in our identification strategy, as discussed ahead. As the Figure shows, there was substantial variation across sectors in tariff reductions. Moreover, tariff cuts were strongly and negatively correlated with pre-liberalization tariff levels: industries with initially higher levels of protection (i.e. tariffs) experienced larger tariff reductions (Kovak, 2013).

The measure of local trade shocks exploits the fact that regions with larger employment shares in industries that experienced greater tariff reductions were more likely to be affected by the trade opening process. Put differently, the unilateral trade liberalization episode is more likely to represent a substantial negative labor demand shock in regions with a larger fraction of its labor force employed in industries that faced larger tariff cuts (relative to regions with a larger fraction of employment in industries less affected). We use tariff data from Kume et al. (2003) to construct the "Regional Tariff Change" (RTC)

[6] A more detailed description of the trade liberalization in Brazil can be found in (Kovak, 2013; Dix-Carneiro and Kovak, 2017).

[7] There were minor changes in tariffs after 1995, which are not relevant compared to the changes that occurred in the 1990-1995 period.

Figure 1: Changes in $\log(1 + \text{tariff})$, 1990-1995



Source: Dix-Carneiro and Kovak (2017).

as proposed by Kovak (2013):

$$RTC_r = \sum_i \beta_{ri} d \ln(1 + \tau_i) \tag{1}$$

where

$$\beta_{ri} = \frac{\frac{\lambda_{ri}}{\theta_i}}{\sum_i \frac{\lambda_{ri}}{\theta_i}}$$

and $\lambda_{ri} = \frac{L_{ri}}{L_r}$ is the fraction of labor allocated to industry i in region r; and θ_i is equal to one minus the wage bill share of industry i. It is worth emphasizing that we use changes in output tariffs to construct the RTC_r. Alternatively, one could use effective rates of protection, which incorporate both input and output tariffs. However, at the level of industry classification used here – which is standard in the literature – changes in input and output tariffs are highly correlated and the regional tariff changes computed using either measure (output tariffs or effective rates of protection) are almost perfectly correlated (Dix-Carneiro et al., 2018).[8]

Since we investigate the effects of trade liberalization on skilled and unskilled workers separately, we also compute different regional tariff shock measures for skilled and unskilled workers, which we denote $RTC_{r,k}$, where k denotes the skill group. For that, we follow a similar approach to that of Autor et al. (2018) and compute weights that are specific to the skill groups, which are given by $\lambda_{rik} = \frac{L_{rik}}{L_{rk}}$.[9] As Figure A.1 in the Appendix shows, the measure of regional trade shock for low-skill workers is highly correlated with the overall measure (RTC_r). This is expected, as low-skill workers correspond to the vast majority of the labor force. The RTC measure for high-skill workers is also strongly correlated with the overall measure, but less than the one for low-skill workers.

2.2 Labor Regulations and Enforcement

In Brazil, the permissible types of labor contracts, their conditions and terms for termination are completely regulated by a labor code based on the civil law system, the *Consolidação das Leis Trabalhistas* (CLT), which dates back to 1943. As part of the labor regulations in Brazil, formal workers are required to hold a booklet issued by the Ministry of Labor that must be signed by the employer and which contains workers' entire formal employment history. Having the labor contract registered in this booklet entitles workers to a series of rights and benefits, such as unemployment insurance, severance payment, a one-month paid vacation and a 50 percent premium for overtime hours.

The labor regulation introduced in 1943 was already quite detailed, extensive and rigid. The new Federal Constitution enacted in 1988 substantially extended the range of labor regulations and workers' benefits, further increasing the regulatory hiring and firing costs.[10] Labor taxes are also quite high in Brazil, with the main components being social security contributions (20%), direct payroll taxes (9%), and contributions to a severance fund (8.5%). Computing the overall labor tax rate as the share of commercial profits (which provides a cross-country comparison), it amounts to 42.1 percent in Brazil, 12.9 percent in Canada and 10 percent in the U.S. Not only the tax rate is high, but there are also substantial compliance costs involved. The time required to pay labor taxes in Brazil is nearly 5 times higher than in the U.S., 491 and 100 hours, respectively (Doing Business, 2007).

[8]In the Appendix, we examine the robustness of our results to different measures of local trade shocks, such as the ratio of imports to production and import penetration coefficient. All of our results remain largely unchanged.

[9]See Dix-Carneiro and Kovak (2015) for a more general discussion within a specific-factors model of regional economies with two types of workers.

[10]See De Barros and Corseuil (2004) for a complete description of the changes introduced by the 1988 Constitution.

Given how cumbersome and costly the labor regulation is, both firms and workers have incentives to either partially comply or avoid it entirely via informal labor contracts. In such an environment where incentives to formalization are arguably weak, enforcement plays a substantial role in determining not only informality levels but labor market outcomes more broadly.

The Ministry of Labor is directly responsible for enforcing labor regulations, but it only inspects registered firms and therefore it does not tackle informal labor in informal firms. Enforcement is implemented in a very decentralized way, both at the state level (with a labor office called *delegacia do trabalho*) and within states through local labor offices called *subdelegacias*. The state level office (*delegacia*) is always located in the state's capital and the local offices (*subdelegacias*) are spread out across municipalities. The number of local offices is a function of the state's size and economic relevance (Almeida and Carneiro, 2012).

Inspectors are allocated to a specific *subdelegacia* and they travel by car to inspect firms. Most inspections are triggered by anonymous reports and inspectors must assess compliance with all the relevant dimensions of the labor code and not only if workers are formally registered or not. There is substantial regional, within-state variation in enforcement intensity (Almeida and Carneiro, 2012). In particular, one of the factors that determines regional variation in enforcement is the relative density of local labor offices, which by its turn determines the travel distances that inspectors face in order to carry out inspections. Thus, these travel distances are a key determinant of the capacity of enforcement across local labor markets.

2.3 Data

Throughout the paper, our main unit of analysis is the micro-region, which is a collection of contiguous municipalities that are economically integrated. The micro-regions are defined by the National Bureau of Statistics (IBGE) and closely reproduce the idea of local economies (similar to Commuting Zones in the US, as in Autor et al. (2013)), which have been extensively used in the recent literature (e.g. Kovak, 2013; Costa et al., 2016; Dix-Carneiro and Kovak, 2017).[11] We use a mapping between municipalities and micro-regions that results in 411 consistent micro-regions between 1980 and 2000. Our analysis focuses on the changes between 1990 and 2000, but we use the 1980 census to control for baseline characteristics.

We use four datasets in our empirical exercise. The first is the Decennial Population

[11] Indeed, the data shows that only 3.2% of workers lived and worked in different micro-regions in 2000 (Dix-Carneiro et al., 2018).

Census, which contains information on individuals' socioeconomic characteristics, as well as labor market outcomes. Particularly important for our exercise, the Census provides information on workers' informality status. We define as informal workers those employees who do not hold a formal contract, which in Brazil is characterized by a "signed work booklet" (as discussed in the previous section). Workers who report being employees are directly asked whether they have a formal contract, which is the information we use to define if an employee is formal or not. Our measure of informality therefore excludes the self employed. We do so because the mechanisms on which we focus here, and in particular our measures of enforcement, refer to employees only. Nevertheless, we also separately analyze the effects on the self-employed, as individuals can respond to worsening labor market conditions by resorting to self employment.

The second data set contains administrative data from the Ministry of Labor related to enforcement activity. This dataset contains yearly information on the number of firms inspected by municipality from 1995 to 2013; number of inspectors responsible for the auditing process in each state of the country; and the location of all labor offices. We add to this administrative data set a crucial piece of information, namely the date of creation of each labor office (i.e. *subdelegacia*). It was necessary to directly call each of the 121 labor offices in Brazil to collect this information. Among those, 92 offices were created prior to 1990 (the start year of the trade opening process), 19 offices were created between 1990 and 2000, and the remaining were created after 2000. The third data set contains the driving distance to the nearest labor office in each municipality, distance to the state's capital and number of inspectors at the state level compiled by Almeida and Carneiro (2012). Finally, we use the *Relação Anual de Informações Sociais* (RAIS), which is an administrative data set collected by the Ministry of Labor that contains the universe of formal firms and workers. We use the RAIS to compute the number of formal establishments in each micro-region.

To obtain a measure of enforcement capacity at the micro-region level, we compute the maximum distance to the nearest labor office within each micro-region.[12] As discussed in the previous section, a greater distance is associated to weaker enforcement capacity in the micro-region. As for the measure of actual enforcement intensity, we compute the ratio between the total number of inspections carried out in 1995-1999 and the total number of formal firms in a given micro-region.

Figure 2 displays the kind of variation we will be exploring throughout our analysis. Panel (a) shows how the regional trade shock, RTC_r, varies across micro-regions in Brazil. Panel (b) displays the regional variation in enforcement intensity (i.e. inspections per

[12]We also use the average distance in our robustness analysis and our results remain largely unchanged.

100 firms), as well as the location of all 92 local labor offices (*subdelegacias*) created prior to 1990 in Brazil.[13] As the figure shows, there is substantial variation in the trade shock, the intensity of enforcement and in the density of local labor offices across micro-regions.

Table 1 provides the descriptive statistics of the variables used in our analysis, which are all constructed at the micro-region level. We define low-skill workers as those with less than completed high school, and high-skill workers as those with at least a high school diploma. Non-employment is defined as the sum of individuals actively looking for jobs (the unemployed) and those out of the labor force. We use this measure (instead of unemployment) to reduce measurement error, as these two states tend to be less distinct in less developed economies (Donovan et al., 2020). The table shows that micro-regions are quite heterogeneous in terms of initial conditions, as most variables have a high level of dispersion. The same is true for the trade shock, which is high on average but also has a very high standard deviation.

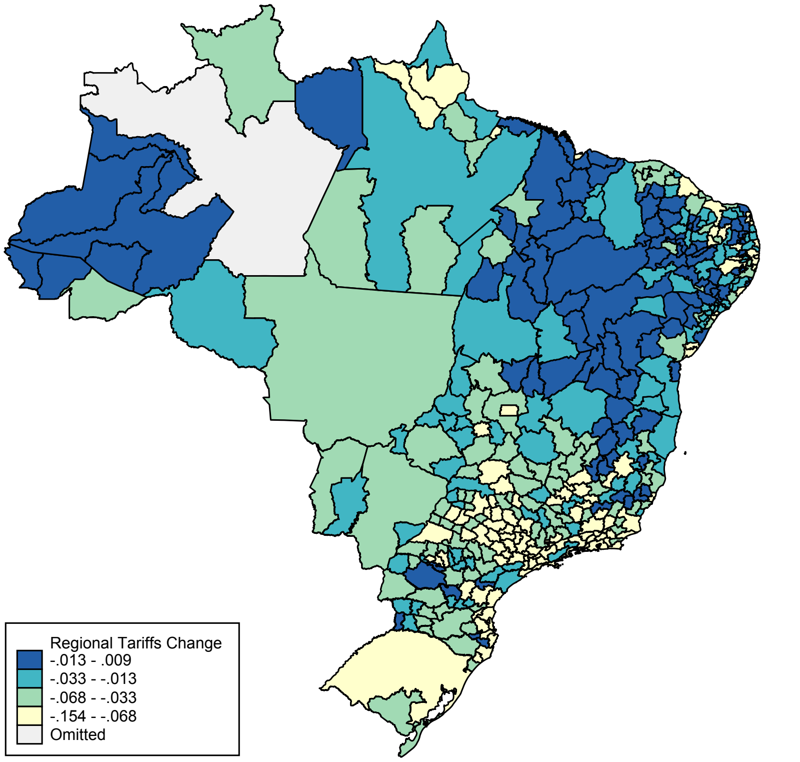
[13] One could also argue that forward-looking policy makers would choose the location of the pre-1990 labor offices based on the expected impact of the tariff reduction policy. However, it is important to notice that a new administration took office in 1990. Hence, all labor offices used in our exercise were created by previous administrations, which could not anticipate the sudden and unexpected change in trade policy.

Table 1: Basic Descriptive Statistics at the Micro-Region Level

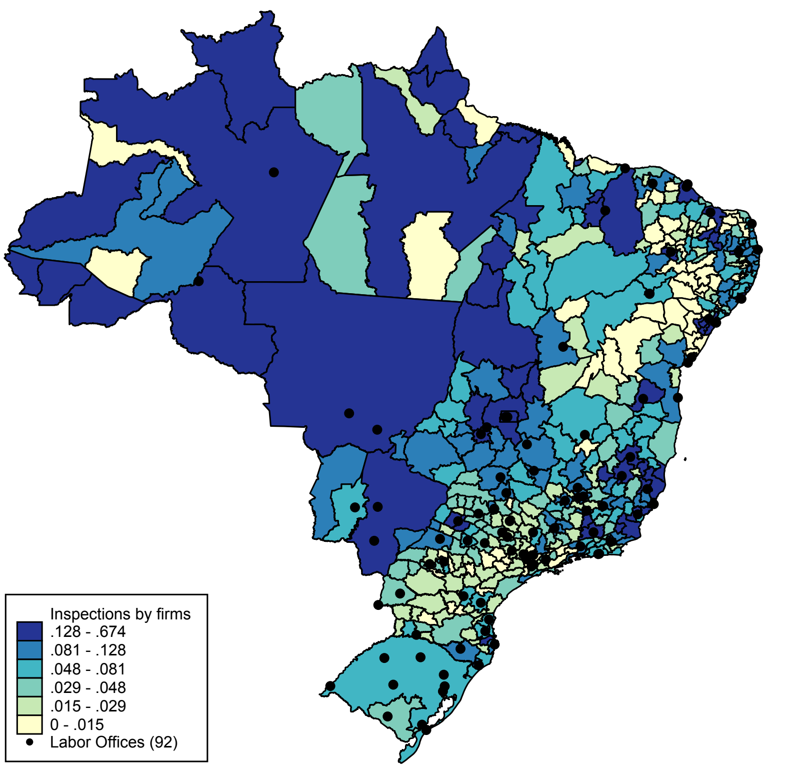
Sample:	All Workers		Low-Skill		High-Skill	
	Mean	SD	Mean	SD	Mean	SD
Non-employment in 1991	0.390	0.045	0.411	0.043	0.221	0.039
Informality in 1991	0.443	0.202	0.464	0.205	0.290	0.170
Self-Employment in 1991	0.182	0.065	0.190	0.068	0.104	0.025
Log-Wages in 1991	0.795	0.420	0.650	0.376	1.605	0.402
RTC_r	0.045	0.040	–	–	–	–
RTC_r Unskilled	–	–	0.043	0.039	–	–
RTC_r Skilled	–	–	–	–	0.091	0.035
Distance L.O. (per 100km)	2.666	1.855	–	–	–	–
Distance to capital	4.096	2.484	–	–	–	–
Number of Inspections per 100 firms	7.430	8.140	–	–	–	–
Share Female in 1991	0.500	0.018	–	–	–	–
Share High-Skill in 1991	0.116	0.060	–	–	–	–
Share Urban Areas in 1991	0.616	0.197	–	–	–	–
Population in 1991	468,949	1,254,580 –	–	–	–	
Number of micro-regions			411			

Notes: We use individual level data and sampling weights from the Demographic Census to compute simple non-employment and informality rates, average wages, share of female, high-skill and urban population at the micro-region level. Distance to L.O., distance to state's capital, number of inspections and the regional trade shocks are calculated as described in the text. Means and standard deviations refer to the distribution of these means at the micro-region level.

Figure 2: Regional Tariff Changes and Enforcement of Labor Regulation across Regions



(a) Distribution of Regional Tariff Changes, RTC_r



(b) Regional distribution of inspections and labor offices

Notes: Map of RCT_r from Dix-Carneiro et al. (2018). Regional distribution of inspections and labor offices obtained using administrative data from the Ministry of Labor (see text).

3 Empirical Strategy

In this section we describe our empirical strategy. However, before proceeding to the discussion of the empirical specifications, it is useful to describe the economic mechanisms that guide our empirical exercise. We do so by using the framework developed in Ulyssea (2018). The interested reader is referred to that paper for a complete discussion of the model, while the reader only interested in the empirical exercise can skip directly to section 3.2.

3.1 Theoretical framework

In the model developed by Ulyssea (2018), firms can exploit two margins of informality. The first is the *extensive margin*, which refers to firms' decision to pay entry fees and register their business or not. The second is the *intensive margin*, which refers to the decision of a formal (registered) firm to hire workers without a formal contract. Firms sort between sectors upon entry based on their expected productivity and the (in)formal sector is comprised by (un)registered firms.

If a firm enters the informal sector, it avoids registration costs and taxes altogether but faces an expected cost associated to informality that is increasing in firm's size. This can be rationalized by the fact that larger firms are more visible to the government and are detected with greater probability.[14] If a firm decides to enter the formal sector, it faces fixed registration costs and must pay revenue and labor taxes. The latter can be avoided by hiring informal workers. However, formal firms that hire informal workers also face a probability of detection, which is increasing in the number of informal workers. Thus, smaller formal firms will hire a larger fraction of their labor force informally and this share decreases with firm size. Since productivity and size are one-to-one in the model, more productive firms (in expectation) self-select into the formal sector and less productive firms enter the informal sector. Similarly, conditional on being formal, less productive firms hire a larger fraction of informal workers.

Firms hire both low- and high-skill workers that are aggregated into a composite labor input through a CES production function. Skill shares may differ across formal and informal firms, and the author's estimates imply that the formal sector is more intensive in high-skill workers. This is consistent with the fact that informality is more prevalent among low-skill workers. The estimation results also indicate that the expected cost of hiring informal workers is lower for low-skill than high-skill individuals, which suggests that higher enforcement would have stronger effects on low-skill workers. The government

[14] This is a common formulation in the literature, see for example de Paula and Scheinkman (2010) and Leal Ordonez (2014), among others.

has two possible enforcement levers: to increase enforcement on the intensive margin of informality, by increasing inspections on formal firms; and to increase enforcement on the extensive margin, by increasing the intensity of inspections on informal firms. In Brazil, these two types of enforcement activities are carried out by different government branches. The enforcement capacity that we examine in this paper refers to enforcement on the intensive margin, which is conducted by the Ministry of Labor.

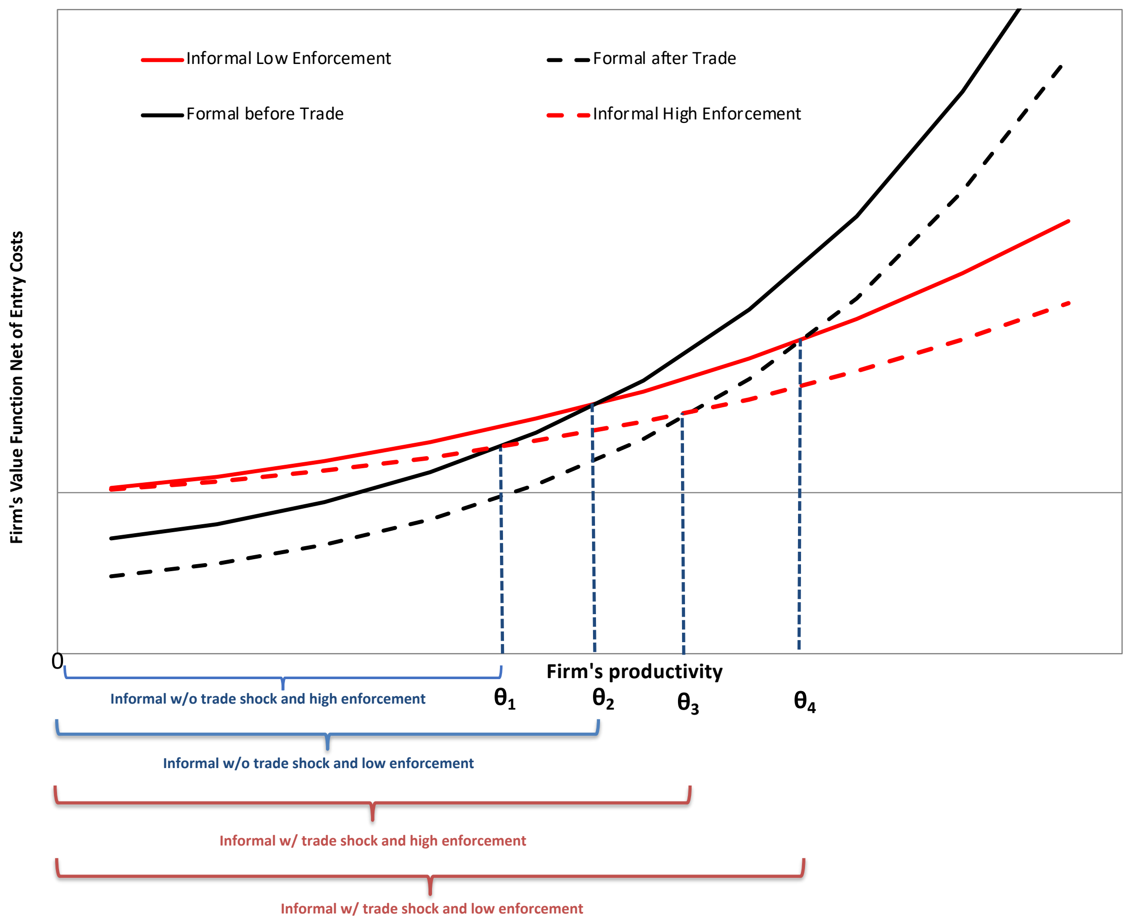
Even though Ulyssea's framework does not include international trade, the unilateral trade opening episode in Brazil can be interpreted through the lenses of the model as a competition shock, which drives down prices for domestic firms (both formal and informal). In order to illustrate the mechanism that we have in mind, we take the estimated model in Ulyssea (2018) and simulate the value functions of being formal before and after the trade opening shock, which is parameterized as a permanent decline in the equilibrium price. Even though in Ulyssea's model prices (i.e. wages) fully adjust in equilibrium, Figure 3 shows the results of a partial equilibrium simulation, where we show firm's payoffs after a one-time price reduction (equivalent to an increase in real wages). For the sake of expositional simplicity, we assume that the negative price shock only affects formal firms, but all that we need to assume is that formal firms are at least as adversely affected as informal ones (but possibly more).

We consider two scenarios for informal firms: low and high enforcement on the extensive margin of informality. We do so because this is the margin that generates greater effects on firms and therefore are easier to visualize in the graph. The exercise would be analogous if we were to consider enforcement on the intensive margin, as they go in the same direction. Figure 3 shows the four corresponding curves.

Consider first the situation prior to the trade shock with the two markets, low and high levels of enforcement (dashed red line, solid black and red lines). In the market with a high level of enforcement, all firms with productivity $\theta < \theta_1$ will optimally choose to be informal as their payoff is higher than the one associated to formality (red dashed line above the black solid line). In the market with a low level of enforcement, firms with productivity $\theta < \theta_2$ will choose to be informal, which shows that for any given distribution of firm productivity the market with lower levels of enforcement will have a larger share of informal firms (and workers), as expected.

When the trade shock hits, the high and low enforcement markets observe an increase in the informality thresholds from θ_1 to θ_3 and θ_2 to θ_4, respectively. However, the impact on informality will be larger in the market with a lower level of enforcement, as more firms can resort to informality to cope with the shock instead of simply exiting. The intuition here is that firms adversely affected can downsize and resort to informality if enforcement is weak (the costs of informality are increasing in firm's size), but this option

Figure 3: Trade Opening under Low and High Enforcement



is not available if enforcement is high. Thus, with weaker enforcement there is a larger increase in informality, but a smaller reduction in the mass of active firms. With stronger enforcement there are lower informality effects, but greater reductions in the mass of active firms. The model does not have unemployment, but the effects on the mass of firms indicate what the effects on unemployment could be, as greater displacement of firms is likely to be associated with higher unemployment. Finally, note that informal firms are more intensive in low-skill workers and formal firms hire a larger fraction of low-skill workers without a formal contract (in contrast to high-skill workers). Hence, the model would predict that these movements towards informality would be stronger for low-skill than high-skill workers.

In sum, this model implies that, in the aftermath of a unilateral trade shock, markets that have lower levels of enforcement would experience a stronger increase in informality, but a lower reduction in the number of plants, including formal ones. Conversely, markets that have higher levels of enforcement would observe a smaller increase in informality, but greater firm exit and potentially larger disemployment effects. The effects on employment and informality would be largely concentrated among low-skill workers.

3.2 Empirical Specification

Our empirical strategy consists of two steps. In the first step, we capture the changes in the outcome of interest at the micro-region level, netting out the influence of individuals' socio-demographic characteristics. More concretely, we run the following regressions at the individual level:

$$Y_{it} = \sum_{r} \gamma_{rt} D_r + \mathbf{x}'_{i,t}\beta_t + \epsilon_{i,t} \tag{2}$$

where i indexes individuals, $t = 1991, 2000$ denotes the year, D_r denotes the set of micro-region dummies, and $\mathbf{x}_{i,t}$ is a vector of individual characteristics that includes age, age squared, schooling, gender and race.

The outcomes considered are wages, a dummy for whether the individual is an informal employee, dummy for self-employed and a dummy for non-employment, which includes both unemployment and out-of-the-labor-force statuses. As discussed in Section 2, the informality dummy considers only those individuals who work as employees in the private sector, and we define as informal workers those without a formal contract (*no signed work booklet*). Since we are analyzing enforcement of labor regulation, which only applies to employees, this informality definition is the most consistent with the goals of our empirical exercise. Nevertheless, we also discuss the effects on self-employment, formal employment and number of formal establishments in Section 4, as these can be important margins of adjustment in the aftermath of the trade shock (Dix-Carneiro and Kovak, 2019) and therefore can shed light on the different forces at play.

The first step thus provides us with a measure of average wages, informality and non-employment rates at the micro-region level. In order to assess the heterogeneous effects across skill groups, we also estimate regression 2 separately for low- and high-skill workers and obtain separate estimates of $\hat{\gamma}$ by skill level.

In the second step, we run regressions in first difference at the micro-region level. The first set of regressions we estimate re-visit the overall labor market impacts of the local trade shock and provide new evidence on the heterogeneity across skill levels. The basic specification is as follows:

$$\Delta\hat{y}_r = \zeta_0 + \zeta_1 RTC_r + \alpha_4 Z_r + \delta_s + u_r \tag{3}$$

where r indexes regions, $\Delta\hat{y}_r \equiv \hat{\gamma}_{r,2000} - \hat{\gamma}_{r,1991}$, Z_r denotes the set of controls used in all regressions and δ_s denotes the state dummies, which absorb differential state-level trends. The vector Z_r includes the controls that are used in all specifications in the paper. It includes the main baseline demographic characteristics of the micro-regions, which could influence the labor market outcomes that we analyze, in particular: the share of women, high-skill individuals, urban population and total population (in log), all measured in 1991.

Our main goal, however, is to investigate how enforcement of labor regulations interacts with the trade shock in shaping local labor market responses. For that, we directly assess how distance to the labor offices might shape the labor market effects from trade. More specifically, we estimate the following regression:

$$\begin{aligned} \Delta\hat{y}_r &= \alpha_0 + \alpha_1 RTC_r + \alpha_2 RTC_r \times Dist_r + \alpha_3 Dist_r + \alpha_4 Z_r \\ &\quad + \alpha_5 Dist_r \times Inspectors_s + \delta_s + \varepsilon_r \end{aligned} \tag{4}$$

where again r denotes the micro-region, δ_s the state dummies, Z_r is the same vector of controls described above (used in all regressions) and $Dist_r$ denotes the maximum distance to the nearest labor office of micro-region r, which is our enforcement capacity variable.[15] We also follow Almeida and Carneiro (2012) and include the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$. The motivation is that, for any given distance, the number of inspectors available at the state level (lowest level of disaggregation available) is an important determinant of enforcement, as it provides a measure of the resources available at the state level.

Since we are using the first-stage estimates as dependent variables, we follow the previous literature and weight the second-stage regressions by the inverse of the standard errors of the first-stage (e.g Kovak, 2013; Dix-Carneiro and Kovak, 2017; Dix-Carneiro et al., 2018).[16] When estimating heterogeneous effects across skill levels, we use the same specification but with the appropriate $\hat{\gamma}$ estimated separately by skill level and the appropriate definition of RTC discussed in the previous section.

[15] We first calculate the distance from the centroid of each municipality to the nearest labor office. The maximum of the distances of the municipalities that belong to the micro-region r is defined as enforcement capacity of micro-region r.

[16] In the Appendix C, we show that our results are robust to not using any weights as well. It is also worth noting that part of the literature uses clustered standard errors at the level of aggregation immediately above micro-regions, which in the Brazilian case would correspond to the meso-region. However, the intra-cluster correlation of our variable of interest – $RTC_r \times Dist_r$ – is very close to zero, which indicates that this clustering is not adequate in our context. Nevertheless, the Appendix C shows that our results are robust to using clustered standard errors at the meso-region level.

Identification: Discussion

A possible critique to our empirical strategy is the absence of random variation in enforcement capacity levels, as distance to the nearest labor office is certainly not randomly distributed across micro-regions. Therefore, one could argue that this variable might be capturing the effect of other characteristics of local economies that are not accounted for in our specifications. We start by noting that, as discussed in Section 2, we restrict ourselves to the labor offices created up until 1990 when constructing the variable $Dist_r$. Thus, we are only using the pre-determined enforcement capacity, which is not responding to the (future) local trade shock and local labor market conditions.

As we estimate the regressions in first difference, our specification accounts for micro-region fixed effects and state-specific trends,[17] as well as differential trends across micro-regions with different initial conditions in terms of demographics and size (we control for the share of women, high-skill individuals, urban population and total population in 1991).

That said, a potential threat to identification could be the existence of reversion to the mean across regions with lower and higher levels of informality and non-employment. If our variable of interest ($RTC_r \times Dist_r$) is simply capturing heterogeneity in initial informality levels, our results would not be identifying the heterogeneous effects of trade across enforcement capacity levels. Instead, they would simply reflect differential trends across high- and low-informality regions. In Section 4.3 we discuss a series of robustness tests, which shows that our results are robust to the inclusion of 1980's informality and non-employment rates in the first-differenced regression (Equation 4). We also examine how our results are affected by the inclusion of other controls that account for potential confounding effects, including: (i) differential trends across more and less remote regions (further away from large urban centers); (ii) local supply of public goods, proxied by local government spending; and (iii) initial level of inequality in the micro-region. As we discuss in Section 4.3, the results remain unchanged and, if anything, become stronger.

Even if one argues that some important determinant of local economic development remains unaccounted for – and that it is being captured by the enforcement capacity measure – one would expect this omitted variable to produce heterogeneous effects on informality and non-employment that go in the same direction. Put differently, it is hard to think of an omitted variable that could lead simultaneously to greater informality and lower disemployment effects as a response to trade liberalization, within the same state and after controlling for the different variables discussed above. Even more so because

[17] The inclusion of state dummies is important because many relevant policies and resources are defined at the state level (e.g. police force and a substantial fraction of health and education expenditures).

this pattern is only observed among low-skill workers, while high-skill workers do not seem to be affected by the level of enforcement in their region.

4 Enforcement of labor regulations and the labor market effects of trade

We start our analysis by re-visiting the basic average results on informality, non-employment and wages found in the literature. Table 2 shows the estimates of regression 3 using as dependent variable the decadal changes in informality (columns 1–3), non-employment (columns 4–6) and wages (columns 7–9). For all three outcomes, we assess the effects on all workers, low-skill and high-skill workers. Importantly, since the variable RTC_r refers essentially to tariff cuts, it is negative for almost all regions. We therefore work with the negative of RTC_r, so that the results are more easily interpretable.

Columns 1, 4 and 7 show the same patterns found in previous studies (e.g Dix-Carneiro and Kovak, 2017; Dix-Carneiro et al., 2018): between 1991 and 2000, regions that were hit harder by the trade opening process experienced an increase in informality and non-employment and a decrease in wages relative to regions less affected. The effects are statistically significant and economically meaningful. To illustrate the magnitude of these effects, consider moving a region from the 10th to the 90th percentile of the distribution of RTC_r, which implies a change of 0.1 log point in the trade shock. The results from Table 2 imply that this micro-region would experience an increase of 4.5 ($0.1 \times 0.451 \times 100$) and 2.1 ($0.1 \times 0.21 \times 100$) percentage points in informality and non-employment rates, respectively. To put these numbers in perspective, they correspond to around 54 and 58 percent of a standard deviation in decadal changes in informality and non-employment rates, respectively.

The novel results in Table 2 come from the analysis of heterogeneous effects across skill levels. As the table shows, almost all negative effects on informality, non-employment and wages come from low-skill workers. High-skill workers show no statistically significant effects on informality and wages, and a much smaller, marginally significant effect on non-employment. Using the same reasoning as above, moving a region from the 10th to the 90th percentile of the distribution of RTC_r would imply an increase of 5.2 ($0.1 \times 0.520 \times 100$) and 2.7 ($0.1 \times 0.267 \times 100$) percentage points in informality and non-employment rates among low-skill workers, respectively. These effects account for 60 and 67 percent of a standard deviation in decadal changes in informality and non-employment rates for low-skill workers, respectively.

Table 2: Basic Effects on Informality, Non-Employment and Wages

	Informality			Non-employment			Wages		
Sample (by workers' skill level):	All	Low	High	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC_r	0.451***			0.206**			-1.062***		
	(0.130)			(0.082)			(0.221)		
RTC-Unskilled$_r$		0.520***			0.267***			-0.930***	
		(0.131)			(0.091)			(0.231)	
RTC-Skilled$_r$			-0.093			0.119*			-0.408
			(0.191)			(0.066)			(0.295)
Observations	411	411	411	411	411	411	411	411	411
R-squared	0.375	0.409	0.315	0.395	0.402	0.319	0.608	0.558	0.588

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions include state fixed effects and the following demographic controls: share of women, high-skill individuals, urban population and log-population in 1991.

4.1 Effects across Enforcement Capacity Levels

We now turn to our main analysis – the heterogeneous effects of trade liberalization across enforcement capacity levels (i.e. distance to the labor office), as described by regression 4. Table 3 shows the results for employment-related outcomes: columns 1-3 show the results for informality and columns 4-6 for non-employment. We estimate regression 4 using all individuals and separately by skill level. The results for all workers indicate that regions with lower enforcement capacity (greater distances to the L.O.) experience stronger effects on informality. The same is not true for non-employment, as the interaction term is small in magnitude and not significant.

However, once we disaggregate the results by skill level, the same striking pattern observed in the basic regression emerges. We find sizable and statistically significant heterogeneous effects for both informality and non-employment among low-skill workers, but none for high-skill workers. For a strong trade opening shock ($RTC_r = 0.1$, as in the previous section), a region with very low enforcement capacity (90th percentile of the distance distribution) would experience an increase of 10 p.p. on informality, but nearly zero effects on non-employment. In contrast, a region with high enforcement capacity (10th percentile of the distance distribution) experiences a lower informality effect – a 3 p.p. increase – but much stronger negative effects on employment, with an increase in non-employment rates of 3.9 p.p.

The intuition for these results is as follows. In regions with stricter enforcement, firms cannot resort to informality and therefore there is more *de facto* labor market rigidity, which leads to greater employment losses. The opposite happens in regions where enforcement capacity is weaker and *de facto* labor market flexibility is higher: there are strong informality effects, but no statistically significant effects on employment. These results therefore indicate that, when enforcement is low, informality acts as an employment buffer in regions more adversely affected by the trade opening episode, but only for low-skill workers.

These stronger informality effects could, in principle, be a combination of two main forces. First, the buffer effect mentioned above, in which formal jobs that would be lost are converted to informal contracts. Second, it could also be the case that formal jobs that would be otherwise retained in a high-enforcement region are converted to informal jobs in low-enforcement regions (i.e. a "switching effect"). While the former could be seen as a potential advantage of informality, as it helps preserving employment, the latter could be seen as detrimental, as it leads to greater losses of formal jobs. To examine this issue we turn to the RAIS data set, which contains the universe of formal establishments and workers. One main advantage is that the RAIS is available since 1988, which allows

Table 3: Effects on Informality and Non-Employment by Enforcement Capacity Level

	Informality			Non-Employment		
Sample (by workers' skill level):	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
RTC_r	0.211			0.348***		
	(0.155)			(0.106)		
$RTC_r \times$ Dist. L.O.$_r$	0.208**			-0.083		
	(0.097)			(0.066)		
RTC-Unskilled$_r$		0.162			0.453***	
		(0.158)			(0.113)	
RTC-Unsk.$_r$ $\times$ Dist. L.O.$_r$		0.326***			-0.148**	
		(0.106)			(0.073)	
RTC-Skilled$_r$			-0.205			0.121
			(0.323)			(0.127)
RTC-Skill.$_r$ $\times$ Dist. L.O.$_r$			0.148			0.004
			(0.199)			(0.089)
Observations	411	411	411	411	411	411
R-squared	0.394	0.434	0.327	0.418	0.428	0.328

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects. The demographic controls are the following: share of women, high-skill individuals, urban population and log-population in 1991.

us to examine pre-trends as well. Given the nature of the data set, we cannot run the first stage regressions at the individual level, so we run the following regressions for each year t:

$$\begin{aligned}\log y_{r,t} - \log y_{r,1991} \quad &= \beta_{0,t} + \beta_{1,t} RTC_r + \beta_{2,t} RTC_r \times Dist_r + \beta_{3,t} Dist_r + \beta_{4,t} X_r \\ &\quad + \beta_{5,t} Dist_r \times Inspectors_r + \delta_{s,t} + \varepsilon_{r,t} \end{aligned} \tag{5}$$

where $y_{r,t}$ represents total formal employment or total number of formal establishments in region r at time $t = 1992, ..., 2000$; we use the same set of controls as in expression 4. For $t = 1988, ..., 1991$, we define the dependent variable as $\log y_{r,1991} - \log y_{r,t}$.

For the sake of simplicity, we only plot the coefficients $\hat{\beta}_{2,t}$ in Figure 4, while the complete results from regression 5 are shown in the Appendix B. The first thing to note is that we find no evidence of pre-trends neither on formal employment, nor on the number of formal establishments before 1991. The heterogeneous effects only become strong and statistically significant from 1994 onwards, after the unilateral trade opening process was concluded. Consistent with the conjecture of the "switching effect", Panel (a) in Figure 4 shows that regions with weaker enforcement experience a stronger reduction in formal employment relative to regions with greater enforcement capacity. However, Panel (b) shows that the greater flexibility introduced by weaker enforcement also leads to greater survival of formal establishments: regions with weaker enforcement capacity (greater distances to labor offices) observe smaller losses in the number of formal plants. This result is consistent with the fact that a large fraction of informal employment is located in formal firms, the so-called intensive margin of informality (Ulyssea, 2018). Moreover, it shows that the intensive margin plays an important role in formal firms' survival in face of an adverse economic shock.

Finally, we examine the effects on wages. As Table 4 shows, we find no statistically significant heterogeneous effects across different enforcement capacity levels. The point estimates for both skill levels are nevertheless large in magnitude and go in the expected direction, i.e. when enforcement is weaker there are greater wage losses. These point estimates are consistent with the idea that the greater *de facto* flexibility provided by informality leads to stronger adjustments in prices (wages) but lower effects on quantities (employment). However, they are not precisely estimated and we cannot reject the hypothesis that the effects are the same across regions with different enforcement capacity levels.[18]

4.2 Self-Employment

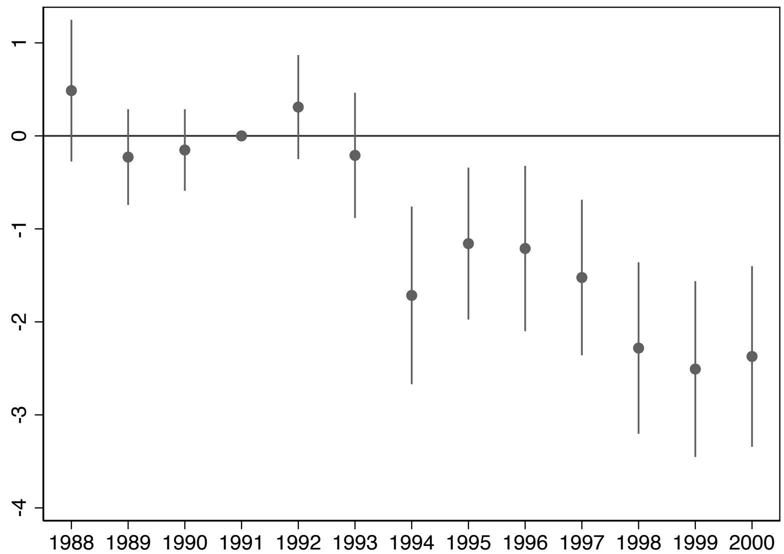
The focus of our analysis lies on informality defined as the share of informal employees. We do so because this definition is the most consistent with our measures of enforcement capacity and intensity. These refer to the enforcement activities conducted by the Ministry of Labor in Brazil, which are targeted at formal firms and their employees.

Another important dimension of informality refers to self-employment. Even though the self-employed are not directly affected by the inspections conducted by the Ministry of Labor, they can be indirectly affected via general equilibrium effects. For example, displaced formal employees might transit to self-employment rather than unemployment. Additionally, our measure of enforcement might be correlated with other dimensions of enforcement, such as inspections on informal firms. Thus, in this section we investigate whether the results observed for labor informality are also observed for self-employment.

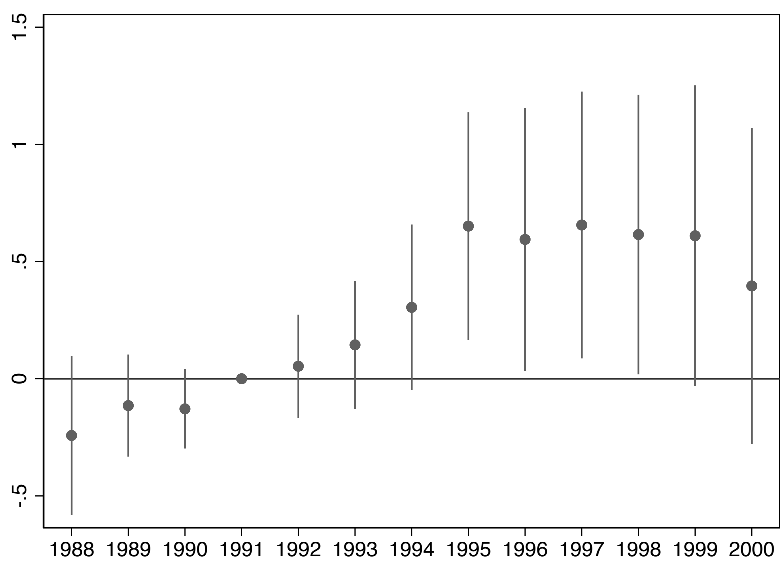
Table 5 shows the results using the specification described in expression 4 and using

[18] However, as we discuss in the robustness analysis (Section 4.3), in some specifications we do find significant heterogeneous effects on low-skill workers' wages.

Figure 4: Effects on the number of formal establishments and workers



(a) Total formal employment



(b) Number of formal establishments

Note: Both panels show estimates of $\hat{\beta}_{2,t}$ from Expression 5.

the share of self-employed among occupied individuals as the outcome of interest. The results show the same patterns observed so far: low-enforcement regions show stronger increases in self-employment relative to regions with higher enforcement capacity (small distances to the L.O.), and these heterogeneous effects are entirely concentrated among

Table 4: Effects on Wages by Enforcement Capacity Level

Sample (by workers' skill level):	All	Low	High
	(1)	(2)	(3)
RTC_r	-0.846***		
	(0.281)		
$RTC_r \times$ Dist. L.O.$_r$	-0.078		
	(0.184)		
RTC-Unskilled$_r$		-0.710**	
		(0.297)	
RTC-Unsk.$_r$ $\times$ Dist. L.O.$_r$		-0.169	
		(0.202)	
RTC-Skilled$_r$			-0.066
			(0.453)
RTC-Skill.$_r$ $\times$ Dist. L.O.$_r$			-0.139
			(0.299)
Observations	411	411	411
R-squared	0.612	0.562	0.596
F-stat (joint significance)	6.271	4.876	0.373
p-value	0.00209	0.00812	0.689

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects. The demographic controls are: share of women, high-skill individuals, urban population and log-population in 1991.

low-skill workers. Considering a strong trade opening shock ($RTC_r = 0.1$), a region with low enforcement capacity (75th percentile of the distance distribution) would experience an increase of 6.5 p.p. on the share of self-employment. This corresponds to 75.7% of a standard deviation in decadal changes in self-employment shares. In contrast, a region with high enforcement capacity (25th percentile of the distance distribution) would experience an increase of 2.6 p.p. on the share of self-employment, or 30% of a standard deviation in decadal changes.[19] These results are consistent with the conjecture that the different margins of enforcement move together and that self-employment is also an important adjustment margin for low-skill formal workers affected by negative labor

[19] Self-employment is strongly counter-cyclical in Latin American countries (e.g. Bosch and Esteban-Pretel, 2012), which is consistent with the substantial magnitudes of our estimated effects in face of a strong negative shock, such as $RTC_r = 0.1$

market shocks.[20]

Table 5: Effects on Self-Employment

Sample (by workers' skill level):	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
RTC_r	0.277**			-0.070		
	(0.121)			(0.137)		
$RTC_r \times$ Dist. L.O.$_r$				0.390***		
				(0.080)		
RTC-Unskilled$_r$		0.332**			-0.029	
		(0.133)			(0.151)	
RTC-Unsk.$_r \times$ Dist. L.O.$_r$					0.425***	
					(0.096)	
RTC-Skilled$_r$			0.103*			0.143
			(0.055)			(0.092)
RTC-Skill.$_r \times$ Dist. L.O.$_r$						-0.032
						(0.052)
Observations	411	411	411	411	411	411
R-squared	0.675	0.726	0.185	0.698	0.744	0.187

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. Regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects. The demographic controls are the following: share of women, high-skill individuals, urban population and log-population in 1991.

4.3 Robustness

In this section, we discuss several different exercises that assess how robust our results are. We start by examining robustness to changes in inference. Table A.2 in Appendix C shows our main results from Tables 3 and 4 with the p-values from our benchmark specification (simple robust standard errors), with clustered standard errors at the meso-region level, and bootstrapped standard errors. Part of the previous literature uses clustered standard errors at the level of aggregation immediately above micro-regions, which in the Brazilian case would correspond to the meso-region (e.g Dix-Carneiro and Kovak, 2015; Dix-Carneiro et al., 2018). As briefly discussed in Section 3, the intra-cluster correlation of our variable of interest – $RTC_r \times Dist_r$ – is very close to zero, which indicates that this

[20]To directly compare the results obtained for informal employees and self-employed, in the Appendix Section C we estimate regression 4 using as outcomes the share of informal employees and self-employed over the working age population. Hence, this exercise uses the same denominator for both employment categories, which allows a direct comparison. The results remain qualitatively unchanged.

clustering is not adequate in our context. We nevertheless report results with clustered standard errors for completeness and comparability with the previous literature. As Table A.2 shows, our results remain as significant as in the benchmark specification. We also bootstrap our entire estimation procedure, which includes the first-stage regressions at the individual level. For that, we use a 10 percent sample of the demographic census and 500 replications. If anything, our results become more significant. Interestingly, the coefficient of the interaction between RTC_r and distance to the labor office in the wage regression for low-skill workers becomes marginally significant (p-value of 0.106).

In Tables A.3-A.11, we re-visit the results from Tables 3 and 4, but including one control at the time and also expanding the set of controls used. These additional controls are the following: informality and unemployment levels in 1980; logarithm of local government spending (per capita) and the Gini coefficient, both measured in 1991; and median driving distance to the state's capital. As discussed in Section 3, the existence of reversion to the mean across regions with lower and higher levels of informality and non-employment could be a potential threat to identification. The inclusion of baseline informality and unemployment rates addresses these concerns.

We obtain the annual local government spending at the municipality level from the Ministry of Finance (*Ministério da Fazenda – Secretaria do Tesouro Nacional*) and then aggregate it at the micro-region level. The inclusion of this variable addresses the possibility that distance to the labor offices can be a proxy for local public goods and infrastructure in general, rather than enforcement capacity *per se*. The initial level of inequality might also be an important determinant of labor market trends and local economic conditions, so we also control for that. Finally, the driving distance to the state's capital tackles the concern of a "remoteness effect", i.e. distance to the labor office could be capturing differential trends across more and less remote regions (further away from large urban centers). These regions are likely to have different levels of development, and can react differently to the trade shocks. In particular, more remote regions might have a higher proportion of low-productivity firms, which are more likely to respond to the trade shocks by increasing informal employment. If this is the case, we would confound this "remoteness effect" with enforcement capacity per se.[21]

As the results in tables A.3–A.8 show, the estimates of the interaction term – trade shock × enforcement capacity levels – generally remain stable or increase in magnitude after we include the basic demographic controls (share of females, high-skill individuals, urban population and log population in 1991). The one exception is the result for non-

[21] It is worth noting that the utilization of this variable implies the loss of 16 micro-regions, for which it is not available. We run the same regressions with the same sample, but without controlling for distance to capital to verify that our results are not driven by this sample variation.

employment among low-skill workers, which becomes slightly weaker and only marginally significant once we include all the additional controls.[22] Nevertheless, RTC_r and the interaction term remain jointly very significant, with a p-value of 0.005. As for wages, tables A.9–A.11 show that including the additional regressors increases precision and the point estimates of the interaction term. The regressions with all additional regressors indicate that regions with weaker enforcement capacity had stronger wage losses, and these effects were entirely concentrated on low-skill workers. This is consistent with the stronger informality effects and lower employment losses also being concentrated among low-skill workers. Finally, tables A.12–A.14 show that the results for self-employment are robust as well.

To further assess the robustness of our conclusions, we estimate a different specification, which is to use the industry-by-micro-region cell as the unit of analysis. In this specification, we move away from our measure of regional tariff change and directly explore the tariff cut variation across industries and the variation in enforcement capacity across regions. The main limitation of this specification is that one cannot analyze the effects on non-employment. More concretely, we estimate the following regression:

$$\Delta y_{r,k} \quad = \beta_0 + \beta_1 \Delta Tariffs_k \times Dist_r + \beta_2 X_{r,k} + \gamma_r + \nu_k + \varepsilon_{r,k}$$

where $\Delta y_{r,k} = y_{r,k,2000} - y_{r,k,1991}$ and $y_{r,k,t}$ is the outcome of interest in region r, industry k and year t; $X_{r,k}$ is the same set of basic demographic controls used in the previous regressions, but computed at the industry-by-micro-region level; and γ_r and ν_k denote region and industry fixed effects, respectively. As Table A.16 shows, we confirm our previous results on informality – strong heterogeneous effects among low-skill workers – but we also find heterogeneous effects for high-skill workers, albeit weaker. We also find significant heterogeneous effects on wages, especially for low-skill workers.

As an additional specification test, in Table A.17 we assess whether the choice of enforcement capacity measure affects our results. Instead of using the maximum distance, we use the average distance to the labor offices as the measure of enforcement. The results remain unchanged.

Finally, in tables A.18 and A.19 we use two alternative measures of local trade shock: (i) the ratio of imports to production and (ii) the import penetration coefficient, respectively. As the tables show, our results on informality remain strong, significant and concentrated among low-skill workers. The effects on unemployment remain strong, but are less precisely estimated and become marginally significant for low-skill workers (p-

[22] Part of this loss in precision could be due to the loss of 16 observations for which there is no information about driving distance to the capital.

values of 0.101 and 0.106). The main disadvantage of these measures is that they are not skill-specific, so they introduce some measurement error.

4.4 Effects across Enforcement Levels: Instrumental Variable Results

To obtain a more direct interpretation in terms of actual enforcement, we estimate an instrumental variable model that has as endogenous regressor the number of inspections per 100 firms in the micro-region. This is the measure of enforcement intensity discussed in Section 2. The second stage regression is analogous to the benchmark specification (Expression 4) and is given by:

$$\Delta\hat{y}_r = \alpha_0 + \alpha_1 RTC_r + \alpha_2 RTC_r \times Enforce_r + \alpha_3 Enforce_r + \alpha_4 Z_r + \delta_s + \nu_r \qquad (6)$$

where all variables remain the same as in 4 and $Enforce_r$ denotes the number of inspections per 100 firms.

The measure of enforcement is clearly endogenous. For example, the government might respond to changes in local labor market conditions by increasing the resources available to enforce the labor regulation in a given market. Conversely, the government might choose to relax enforcement of labor regulations in face of a negative labor demand shock. Moreover, this variable can be subject to substantial measurement error, which also highlights the need for an instrument. We use $Dist_r$ and $RTC_r \times Dist_r$ as instruments for $Enforce_r$ and $RTC_r \times Enforce_r$. We use as an additional instrument the distance ($Dist_r$) interacted with the number of inspectors at the state level. This term captures the fact that, for a given distance, states with more inspectors will have more effective enforcement (conditional on state-specific trends captured by state dummies). We estimate 6 using the limited information maximum likelihood estimator, as it is known to have better small sample properties than two-stage least square in the presence of weak instruments.[23]

Table 6 shows the first set of IV estimates pooling all workers together, while Table 7 reports the results for low- and high-skill workers separately (all first-stage results can be found in the appendix Table A.20). We focus on informality and non-employment only, as the results for wages confirm those discussed in the previous section (i.e. we find no significant effects). Table 6 shows the expected patterns: regions with more enforcement (i.e. more inspections) show lower informality effects and higher unemployment effects,

[23] We also compare the LIML estimates with those obtained with a just-identified two-stage least squares and find very similar and statistically identical results. These are available upon request.

although the former are not statistically significant (column 2).

Once again, the results by skill levels show a striking pattern, as all effects are concentrated on low-skill workers (see Table 7). In regions with more enforcement there are lower informality effects and much higher disemployment effects, but only among low-skill workers. Interestingly, contrasting the OLS and IV estimates indicates that there is an attenuation bias (both in Tables 6 and 7). That is, the OLS estimates tend to underestimate the negative effect on informality and the positive effect on unemployment of higher levels of enforcement. This is consistent with the presence of measurement error in the number of inspections, but also with the government responding to this economic shock by weakening enforcement in regions that are hit harder by the trade opening.

We report the relevant diagnostic statistics in the bottom of both tables. We report the Kleibergen and Paap (2006) test statistics for under-identification and weak instruments. The Kleibergen-Paap LM statistic indicates that we can strongly reject the null that the model is underidentified in all of our results. Similarly, the Kleibergen and Paap (2006) heteroskedasticity-robust Wald rk F-statistic shows that we can reject the null that the instruments are weak for both outcomes for the pooled and low-skill workers sample, but not for high-skill workers.[24] Nevertheless, we use the Anderson and Rubin (1949) weak-instrument robust Wald test for assessing the joint significance of our endogenous regressors ($Enforcement_r$ and $Enforcement_r \times RTC_r$). The results indicate that the endogenous regressors are strongly significant both for non-employment and informality, with the exception of the results for non-employment of high-skill workers.

[24] The critical values for the Kleibergen-Paap test have not yet been tabulated in the literature, but the common practice is to use the critical values of Stock and Yogo (2005).

Table 6: Effects on Informality and Non-Employment: IV Estimation

	Informality		Non-employment	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
RTC_r	0.259	0.784***	0.238**	-0.275
	(0.157)	(0.254)	(0.095)	(0.188)
$RTC_r \times$ Inspections	0.014*	-0.036	0.007	0.068***
	(0.008)	(0.031)	(0.005)	(0.024)
Observations	411	411	411	411
R-squared	0.381	0.261	0.413	0.185
Kleibergen-Paap rk LM statistic	–	12.44	–	14.02
P-value	–	0.002	–	0.001
Kleibergen-Paap rk Wald F statistic	–	6.692	–	6.818
Anderson-Rubin ξ^2 test (robust inference)	–	14.23	–	15.62
P-value	–	0.003	–	0.001

Notes: Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. Estimates obtained using the limited information maximum likelihood estimator. All regressions control for state fixed effects and the following demographic controls: share of women, high-skill individuals, urban population and log-population in 1991.

Table 7: Effects on Informality and Non-Employment by Workers' Skill Level: IV Estimation

	Low-skill				High-skill			
	Informality		Non-Employment		Informality		Non-Employment	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RTC_r-Unskilled	0.293*	1.292***	0.324***	-0.307				
	(0.161)	(0.314)	(0.107)	(0.210)				
RTC_r-Unsk.$\times$ Inspections	0.010	-0.087**	0.006	0.078***				
	(0.008)	(0.038)	(0.005)	(0.026)				
RTC_r-Skilled					-0.103	-0.829	0.064	-0.332
					(0.249)	(0.799)	(0.083)	(0.304)
RTC_r-Skilled$\times$ Inspections					0.011	0.080	0.008	0.065
					(0.022)	(0.133)	(0.009)	(0.050)
Observations	411	411	411	411	411	411	411	411
R-squared	0.418	0.164	0.417	0.152	0.327	0.175	0.322	0.175
Kleibergen-Paap rk LM statistic	–	11.89	–	12.78	–	6.71	–	5.56
P-value	–	0.003	–	0.002	–	0.035	–	0.062
Kleibergen-Paap rk Wald F statistic	–	6.945	–	6.488	–	2.584	–	2.070
Anderson-Rubin ξ^2 test (robust inference)	–	20.08	–	16.42	–	10.62	–	5.633
P-value	–	0.000	–	0.001	–	0.014	–	0.131

Notes: Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. Estimates obtained using the limited information maximum likelihood estimator. All regressions control for state fixed effects and the following demographic controls: share of women, high-skill individuals, urban population and log-population in 1991.

We provide a graphic representation of these results by plotting the effects of the regional trade shock on informality and non-employment as a function of different enforcement levels. More concretely, we plot the following parameter:

$$\hat{\zeta}(Enforcement) = \hat{\alpha}_1 + \hat{\alpha}_2 \times Enforcement \qquad (7)$$

which gives the effect of the regional trade shock for a given enforcement level.

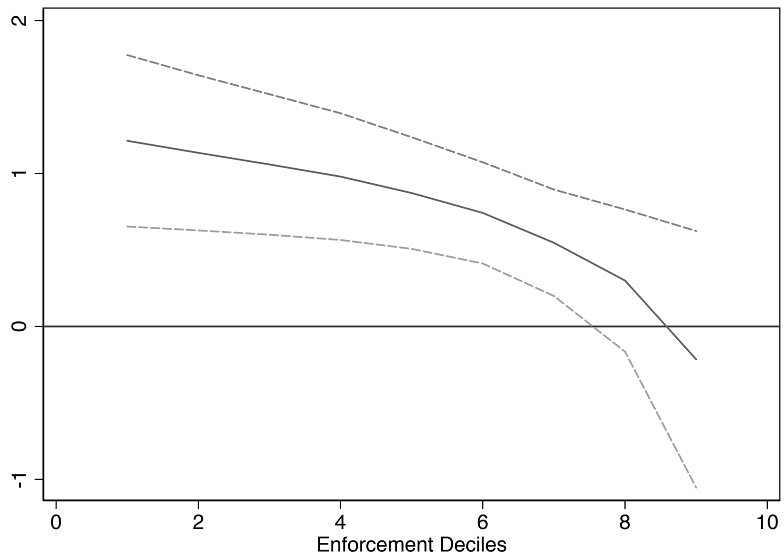
To make the results easily interpretable, we plot $\hat{\zeta}(Enforcement)$ for each decile of the enforcement distribution, i.e. each decile of the distribution of inspections per 100 firms across micro-regions. Since the previous results show that all the effects come from low-skill workers, we only show the plots for this group.

Figure 5 shows a very clear pattern: in regions with low levels of enforcement, the local trade shock produces strong informality effects but no disemployment effects; in regions with high levels of enforcement, the trade shock does not lead to any increase in informality but has very strong disemployment effects. In order to assess the magnitude of these heterogeneous effects, we again consider a high intensity local trade shock ($RTC_r = -0.1$). A region in the first decile of enforcement – with 0.9 inspections per 100 firms – would experience an increase of 12.1 percentage points in informality but no disemployment effects. A region in 90th percentile of enforcement – with 17.2 inspections per 100 firms – would experience no informality effects but an increase of 10.3 percentage points in non-employment rates. Considering the results from Table 2, on average the regional trade shock is associated to an increase of 5.2 and 2.7 percentage points in informality and non-employment among low-skill workers, respectively. Therefore, the lack or strength of enforcement can lead to labor market responses in either informality or non-employment that are substantially larger than the effects observed on average (without accounting for heterogeneity).

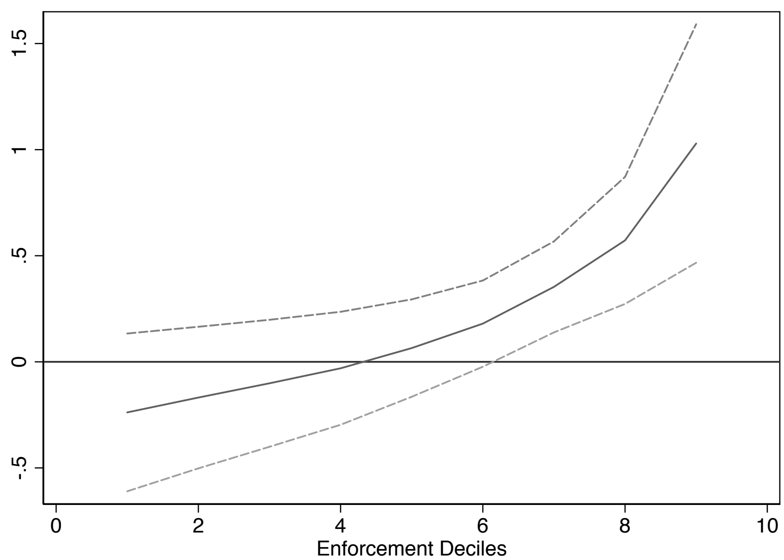
These results are very much in line with the framework discussed in Section 3. In regions where there is greater enforcement, firms are more likely to be detected by the government and therefore the expected cost of informality is higher. Hence, firms are less likely to respond to a negative shock by hiring a greater share of their workers informally. This inability to resort to informality as a coping strategy implies that more firms either fire workers or exit the industry altogether, which implies greater employment losses. The opposite would be observed in a market with low levels of enforcement.

It is important to emphasize that our measure of enforcement is mostly relevant for the intensive margin of informality, as the labor offices only audit formal firms. Nevertheless, the results from Section 4.2 suggest that our measures of enforcement capacity and intensity are likely to be positively correlated with other margins of enforcement (such

Figure 5: Informality and Non-employment Effects Across Enforcement Deciles – low-skill Workers



(a) Informality



(b) Non-employment

as enforcement of tax regulation), which would also discourage the extensive margin of informality. If this is the case, then our empirical estimates are capturing the importance of both margins of informality for firms and workers to cope with adverse economic shocks.

5 Final Remarks

This paper investigates how, and to what extent, enforcement of labor regulations shapes the labor market effects of trade. We do so in the context of Brazil, a country that underwent a major unilateral trade liberalization episode in early 1990s, and which is characterized by burdensome labor regulations that are imperfectly enforced. We exploit variation across regions in both the intensity of the trade shock and enforcement of labor regulations to assess whether different levels of enforcement lead to heterogeneous effects on employment, informality, wages and number of surviving formal plants across local labor markets.

The main results of the paper show that, in the 10 years after the trade opening, regions with stricter enforcement observed: (i) substantially lower informality effects; (ii) much larger disemployment effects; (iii) lower reductions on formal employment; and (iv) greater reductions in the number of formal plants. Regions with weaker enforcement observed symmetric effects. All the effects are concentrated on low-skill workers, while the effects on high-skill workers do not seem to vary across high- and low-enforcement regions. Our results thus indicate that greater *de facto* labor market flexibility introduced by informality allows both formal firms and low-skill workers to cope better with adverse labor market shocks.

We believe these results highlight the importance of labor market rigidities introduced by burdensome and costly regulations, and how they can amplify employment losses in face of negative shocks. Even though our empirical setting is a mid-income country with high informality, our results can be relevant for developed countries as well. This is particularly true given the rise of the "Gig Economy" and the increasingly common co-existence between more stable, permanent labor contracts and very flexible, largely unregulated employment relations in more developed economies. Even though the co-existence of these different types of contracts generates non-trivial distributional issues, our results indicate that having the option to rely on more flexible contracts may help prevent greater job losses and firm exit in face of a negative economic shock.

The welfare implications of our results, however, remain as an open question. On the one hand, our findings indicate that the *de facto* flexibility introduced by informality can prevent greater employment losses and formal firm closures in face of a negative demand shock. On the other hand, the literature has extensively shown that, on average, informal workers have lower earnings and worse working conditions than their (observationally equivalent) formal counterparts. Moreover, results on wages suggest that workers in regions with weaker enforcement might experience wage losses, which potentially introduce additional distributional concerns (although these results are not precisely estimated).

Finally, workers might attribute some intrinsic value to formal jobs and their benefits. Hence, there are important trade-offs at the individual level, and the net welfare effects are unclear *a priori*.

Finally, even if informality acts as an employment buffer in the mid-run – in which case one could argue it is a second-best relative to a counterfactual scenario with perfect enforcement and no informality – it remains unclear what the long-run consequences are. Hysteresis can lead to persistent informality effects even after the dissipation of the economic shock. If informal jobs are indeed of lower quality, and given that informality can lead to substantial misallocation of resources and worse aggregate outcomes (see Ulyssea, 2020, for a review), workers in low-enforcement regions can be worse off in the long run.

References

Adhvaryu, A., A. V. Chari, and S. Sharma (2013). Firing costs and flexibility: evidence from firms' employment responses to shocks in india. *Review of Economics and Statistics 95*(3), 725–740.

Almeida, R. and P. Carneiro (2012). Enforcement of labor regulation and informality. *American Economic Journal: Applied Economics 4*(3), 64–89.

Almeida, R. K. and J. P. Poole (2017). Trade and labor reallocation with heterogeneous enforcement of labor regulations. *Journal of Development Economics 126*, 154–166.

Anderson, T. W. and H. Rubin (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of Mathematical Statistics*, 46–63.

Autor, D., D. Dorn, and G. Hanson (2013, October). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review 103*(6), 2121–68.

Autor, D., D. Dorn, G. Hanson, et al. (2018). When work disappears: Manufacturing decline and the falling marriage-market value of young men. *American Economic Review: Insights*.

Barros, R. P. d. and C. H. Corseuil (2004). The impact of regulations on brazilian labor market performance. In J. Heckman and C. Pages (Eds.), *Law and Employment: Lessons from Latin America and the Caribbean*, pp. 273–350. University of Chicago Press.

Besley, T. and R. Burgess (2004, February). Can labor regulation hinder economic performance? evidence from india. *The Quarterly Journal of Economics 119*(1), 91–134.

Boeri, T. and P. Garibaldi (2005). Shadow sorting. In C. Pissarides and J. Frenkel (Eds.), *NBER Macroeconomics Annual*. MIT Press.

Bosch, M. and J. Esteban-Pretel (2012). Job creation and job destruction in the presence of informal markets. *Journal of Development Economics 98*(2), 270–286.

Bosch, M., E. Goñi-Pacchioni, and W. Maloney (2012). Trade liberalization, labor reforms and formal–informal employment dynamics. *Labour Economics 19*(5), 653–667.

Botero, J. C., S. Djankov, R. L. Porta, F. Lopez-de Silanes, and A. Shleifer (2004). The regulation of labor. *The Quarterly Journal of Economics 119*(4), 1339–1382.

Costa, F., J. Garred, and J. P. Pessoa (2016). Winners and losers from a commodities-for-manufactures trade boom. *Journal of International Economics 102*, 50–69.

De Barros, R. P. and C. H. Corseuil (2004). The impact of regulations on brazilian labor market performance. In *Law and Employment: Lessons from Latin America and the Caribbean*, pp. 273–350. University of Chicago Press.

de Paula, A. and J. A. Scheinkman (2010). Value-added taxes, chain effects, and informality. *American Economic Journal: Macroeconomics 2*(4), 195–221.

Dix-Carneiro, R. and B. K. Kovak (2015). Trade liberalization and the skill premium: A local labor markets approach. *American Economic Review 105*(5), 551–57.

Dix-Carneiro, R. and B. K. Kovak (2017). Trade liberalization and regional dynamics. *American Economic Review 107*(10), 2908–46.

Dix-Carneiro, R. and B. K. Kovak (2019). Margins of labor market adjustment to trade. *Journal of International Economics 117*, 125–142.

Dix-Carneiro, R., R. Soares, and G. Ulyssea (2018). Economic shocks and crime: evidence from the brazilian trade liberalization. *American Economic Journal: Applied Economics 10*(4), 158–95.

Doing Business (2007). Paying taxes - the global picture. The World Bank: Washington, DC.

Donovan, K., W. Lu, and T. Schoellman (2020). Labor market dynamics and development. Mimeo.

Freeman, R. B. (2010). Labor regulations, unions, and social protection in developing countries: market distortions or efficient institutions? In *Handbook of development economics*, Volume 5, pp. 4657–4702.

Goldberg, P. K. and N. Pavcnik (2003). The response of the informal sector to trade liberalization. *Journal of Development Economics 72*(2), 463–496.

Goldberg, P. K. and N. Pavcnik (2007). Distributional effects of globalization in developing countries. *Journal of economic literature 45*(1), 39–82.

Hakobyan, S. and J. McLaren (2016). Looking for local labor market effects of nafta. *Review of Economics and Statistics 98*(4), 728–741.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of econometrics 133*(1), 97–126.

Kovak, B. K. (2013). Regional effects of trade reform: What is the correct measure of liberalization? *The American Economic Review 103*(5), 1960–1976.

Kume, H., G. Piani, and C. Souza (2003). A politica brasileira de importacao no periodo 1987-1998: Descrição e avaliaçcao. In C. H. Corseuil and H. Kume (Eds.), *A Abertura Comercial Brasileira Nos Anos 1990: Impacto Sobre Emprego e Salario*, pp. 9–37. IPEA.

Leal Ordonez, J. C. (2014). Tax collection, the informal sector, and productivity. *Review of Economic Dynamics 17*(2), 262–286.

Meghir, C., R. Narita, and J.-M. Robin (2015). Wages and informality in developing countries. *American Economic Review 105*(4), 1509–46.

Menezes-Filho, N. A. and M.-A. Muendler (2011). Labor reallocation in response to trade reform. NBER Working Paper No. 17372.

Nickell, S. and R. Layard (1999). Labor market institutions and economic performance. Volume 3, Part C of *Handbook of Labor Economics*, pp. 3029 – 3084. Elsevier.

Nickell, S., L. Nunziata, and W. Ochel (2004). Unemployment in the oecd since the 1960s. what do we know? *The Economic Journal 115*(500), 1–27.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear iv regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80.

Tella, R. D. and R. MacCulloch (2005). The consequences of labor market flexibility: Panel evidence based on survey data. *European Economic Review 49*(5), 1225 – 1259.

Topalova, P. (2010). Factor immobility and regional impacts of trade liberalization: Evidence on poverty from India. *American Economic Journal: Applied Economics 2*(4), 1–41.

Ulyssea, G. (2010, January). Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics 91*, 87–99.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from brazil. *American Economic Review 108*(8), 2015–47.

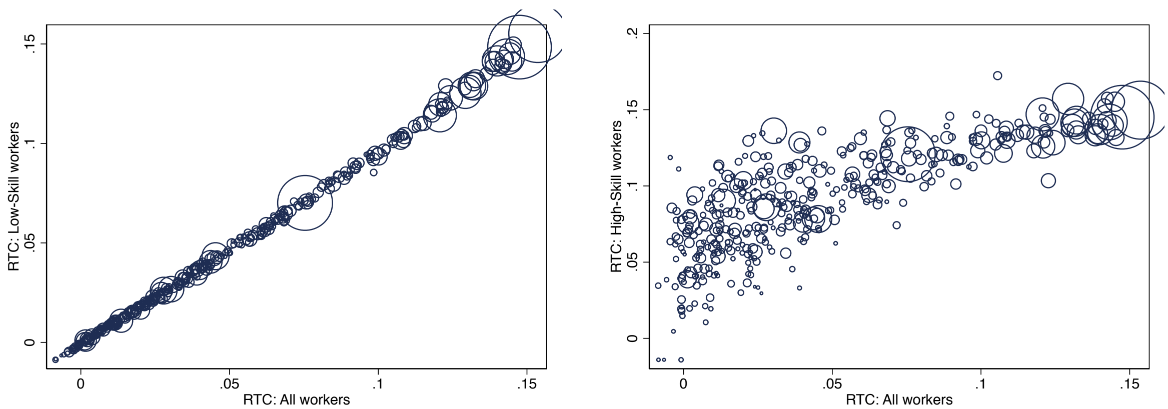
Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics 12*.

Appendix

A Additional descriptive statistics

Figure A.1 shows the relationship between the overall measure of Regional Trade Shock (RTC_r) and the measures for low- and high-skill workers, $RTC_{r,\text{low-skill}}$ and $RTC_{r,\text{high-skill}}$, respectively. These measures are computed following expression 1, with skill-specific weights given by $\lambda_{rik} = \frac{L_{rik}}{L_{rk}}$, $k = low, high$.

Figure A.1: Regional Tariff Changes



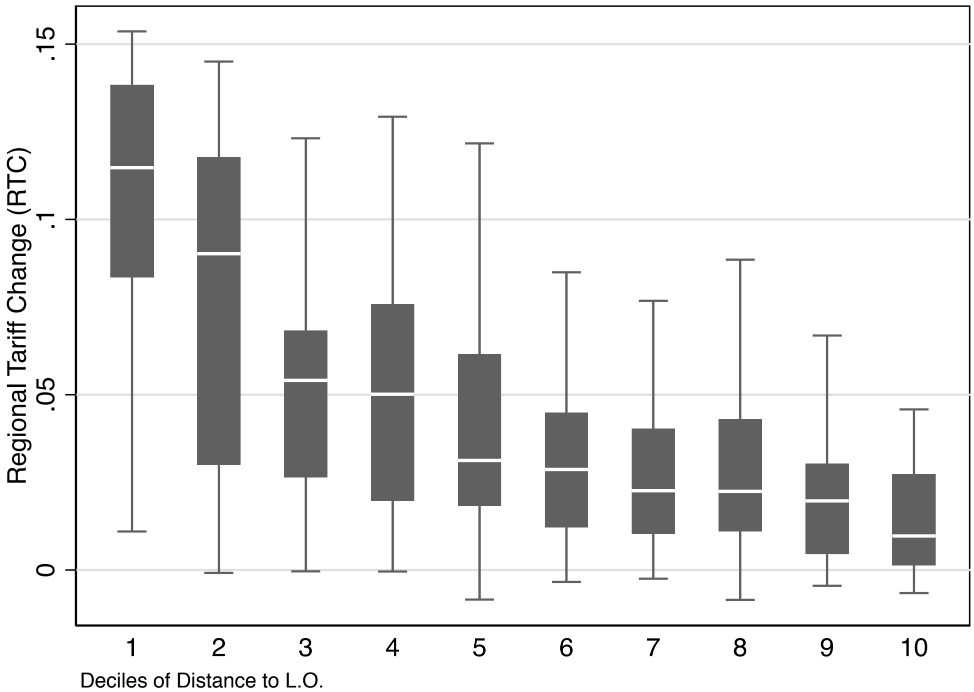
(a) $RTC_{r,\text{low-skill}}$ vs. RTC_r

(b) $RTC_{r,\text{high-skill}}$ vs. RTC_r

Notes: Weights are given by average population at each micro-region in 1991 and 2000.

To assess whether there is enough variation in the regional trade shock, RTC_r over the support of our measure of enforcement capacity, $Dist_r$, Figure A.2 shows a box plot of RTC_r for each decile of the distribution of $Dist_r$.

Figure A.2: Distribution of RTC_r for each decile of the distribution of distance to the labor offices ($Dist_r$)



B Results using the RAIS data

Here we report the complete results using the RAIS data, which contain the universe of formal establishments and workers in Brazil. As discussed in the text, the estimated regression is the following:

$$\begin{aligned}\log y_{r,t} - \log y_{r,1991} &= \beta_{0,t} + \beta_{1,t} RTC_r + \beta_{2,t} RTC_r \times Dist_r + \beta_{3,t} Dist_r + \beta_{4,t} X_r \\ &\quad + \beta_{5,t} Dist_r \times Inspectors_r + \delta_{s,t} + \varepsilon_{r,t}\end{aligned}$$

where $y_{r,t}$ represents total formal employment or total number of formal establishments in region r at time $t = 1992, ..., 2000$; we use the same set of controls as in expression 4. For $t = 1988, ..., 1991$, we define the dependent variable as $\log y_{r,1991} - \log y_{r,t}$.

Table A.1: Effects on formal establishments and workers

	1988	1989	1990	1992	1993	1994	1995	1996	1997	1998	1999	2000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
					Panel A: Number of formal establishments							
RTC_r	1.258***	0.776***	0.485***	-0.605***	-0.432*	-0.288	-1.471***	-1.256**	-1.255**	-0.945	-0.681	-0.318
	(0.325)	(0.223)	(0.142)	(0.194)	(0.248)	(0.320)	(0.505)	(0.554)	(0.593)	(0.652)	(0.681)	(0.714)
$RTC_r \times$ Dist. L.O.	-0.242	-0.115	-0.129	0.053	0.144	0.305*	0.651***	0.594**	0.656**	0.615**	0.610*	0.396
	(0.173)	(0.111)	(0.086)	(0.112)	(0.139)	(0.180)	(0.248)	(0.286)	(0.290)	(0.304)	(0.327)	(0.344)
Observations	411	411	411	411	411	411	411	411	411	411	411	411
R-squared	0.415	0.556	0.742	0.953	0.669	0.677	0.648	0.669	0.730	0.755	0.769	0.787
					Panel B: Number of formal workers							
RTC_r	0.032	-0.051	0.033	-0.139	-0.211	0.470	-0.484	-0.449	0.068	0.898	1.847*	1.771*
	(0.566)	(0.613)	(0.414)	(0.493)	(0.554)	(0.775)	(0.764)	(0.767)	(0.823)	(0.907)	(0.959)	(0.975)
$RTC_r \times$ Dist. L.O.	0.487	-0.228	-0.152	0.310	-0.210	-1.715***	-1.158***	-1.211***	-1.523***	-2.281***	-2.507***	-2.372***
	(0.389)	(0.263)	(0.224)	(0.285)	(0.344)	(0.487)	(0.416)	(0.453)	(0.427)	(0.470)	(0.482)	(0.495)
Observations	411	411	411	406	411	411	411	411	411	411	411	411
R-squared	0.272	0.427	0.216	0.283	0.268	0.462	0.439	0.485	0.525	0.612	0.644	0.660

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 5, which includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects. The demographic controls are: share of women, high-skill individuals, urban population and log-population in 1991.

C Robustness analysis

This section reports the robustness analysis. Table A.2 shows our main results from Tables 3 and 4 with the p-values from the benchmark specification (simple robust standard errors), with clustered standard errors at the meso-region level, and bootstrapped standard errors. Tables A.3-A.11 re-visit the results from Tables 3 and 4, but including one control at the time and also expanding the set of controls used. These additional controls are the following: informality and unemployment levels in 1980; logarithm of local government spending (per capita) and the Gini coefficient, both measured in 1991; and median driving distance to the state's capital. Tables A.12–A.14 contain the same exercise, but for the effects on self-employment.

To obtain directly comparable measures for informal employment and self-employed, we have computed both categories as the share of the working age population (18 to 64 years old, as defined in the paper) in each micro-region r in 1991. By doing this, we hold the denominator fixed and hence it is not affected by the trade shock. Additionally, we have also computed the fraction of employees (formal and informal) in the working age population in each micro-region r in 1991. The number of employees is the denominator used in our definition of informality. Thus, with this exercise we can separately assess the impact of the shock on the numerator (informal employees) and denominator (formal and informal employees) of our main measure of informality.

Table A.16 contains the results of the specification that uses the industry-by-micro-region cell as the unit of analysis. In this case, we run the following regression:

$$\Delta y_{r,k} = \beta_0 + \beta_1 \Delta Tariffs_k \times Dist_r + \beta_2 X_{r,k} + \gamma_r + \nu_k + \varepsilon_{r,k}$$

where $\Delta y_{r,k} = y_{r,k,2000} - y_{r,k,1991}$ and $y_{r,k,t}$ is the share of informal workers in region r, industry k and year t; $X_{r,k}$ is the same set of basic demographic controls used in the previous regressions, but computed at the industry-by-micro-region level; and γ_r and ν_k denote region and industry fixed effects, respectively.

We also examine how sensitive our results are to changing the way we compute the enforcement capacity variable. In Table A.17, we use the average distance to the labor offices in the micro-regions – instead of the maximum distance – as our measure of enforcement capacity. Again, our results remain unaltered or become a little stronger.

Finally, Tables A.18 and A.19 show the results using the ratio of imports to GDP and import penetration coefficient $(M/(Y+M-X))$ as measures of trade shock, respectively. We report results using the benchmark model given by Expression 4.

Table A.2: Robustness analysis: Inference

	Informality			Non-employment			Wages		
Sample (by workers' skill level):	All	Low	High	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC_r	0.211			0.348			-0.846		
	(0.173)			(0.001)			(0.003)		
	[0.283]			[0.003]			[0.016]		
	{0.646}			{0.002}			{0.002}		
$RTC_r \times$ Dist. L.O.	0.208			-0.083			-0.078		
	(0.032)			(0.212)			(0.670)		
	[0.034]			[0.193]			[0.716]		
	{0.080}			{0.000}			{0.034}		
RTC_r-Unskilled		0.162			0.453			-0.710	
		(0.305)			(0.000)			(0.017)	
		[0.357]			[0.000]			[0.051]	
		{0.578}			{0.002}			{0.002}	
RTC_r-Unsk.$\times$ Dist. L.O.		0.326			-0.148			-0.169	
		(0.002)			(0.044)			(0.403)	
		[0.007]			[0.033]			[0.474]	
		{0.066}			{0.000}			{0.106}	
RTC_r-Skilled			-0.205			0.121			-0.066
			(0.526)			(0.340)			(0.885)
			[0.586]			[0.357]			[0.895]
			{0.368}			{0.186}			{0.448
RTC_r-Skilled$\times$ Dist. L.O.			0.148			0.004			-0.139
			(0.458)			(0.965)			(0.644)
			[0.492]			[0.970]			[0.736]
			{0.526}			{0.282}			{0.400}
Observations	411	411	411	411	411	411	411	411	411
R-squared	0.394	0.434	0.327	0.418	0.428	0.328	0.612	0.562	0.596

Notes: For each coefficient, we report p-values in the following order: baseline (parentheses); cluster at the meso-region level (squared brackets); bootstrap (curled brackets). Regressions follow expression 4, which includes $Inspectors_s \times Dist_r$, distance in levels ($Dist_r$) and state fixed effects. Demographic controls are: share of women, high-skill individuals, urban population and log-population in 1991.

Table A.3: Effects on Informality: All workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC_r	0.094	0.029	0.215*	0.396**	0.211	-0.387**	-0.323**	-0.485***	-0.457***
	(0.106)	(0.117)	(0.126)	(0.155)	(0.155)	(0.153)	(0.163)	(0.162)	(0.170)
$\text{RTC}_r \times \text{Dist. L.O.}_r$	0.098	0.107	0.153	0.149	0.208**	0.161*	0.156*	0.205**	0.227***
	(0.096)	(0.096)	(0.103)	(0.104)	(0.097)	(0.089)	(0.089)	(0.084)	(0.082)
Share female_{1991}		0.285	0.718**	0.731**	0.445	0.053	0.023	-0.137	-0.463
		(0.290)	(0.350)	(0.348)	(0.366)	(0.341)	(0.343)	(0.351)	(0.411)
Share high-skill_{1991}			-0.237***	-0.187**	-0.253***	-0.253***	-0.290***	-0.430***	-0.440***
			(0.085)	(0.084)	(0.081)	(0.074)	(0.083)	(0.085)	(0.083)
Share urban_{1991}				-0.070**	-0.067**	-0.054**	-0.054**	0.007	0.007
				(0.031)	(0.030)	(0.027)	(0.027)	(0.030)	(0.032)
$\log(population_{1991})$					0.013***	0.011***	0.011***	0.008**	0.008**
					(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
$\text{Informality}_{1980}$						0.243***	0.239***	0.255***	0.254***
						(0.029)	(0.029)	(0.029)	(0.030)
$Gini_{1991}$							0.050	0.016	0.032
							(0.044)	(0.043)	(0.043)
$\log(\text{Govt. Spending}_{1991})$								-0.029***	-0.030***
								(0.007)	(0.007)
Distance to Capital									-0.001
									(0.001)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.352	0.354	0.365	0.372	0.394	0.492	0.494	0.511	0.527

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.4: Effects on Informality: Low-skill workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC-Unskilled$_r$	0.048	-0.044	0.178	0.386**	0.162	-0.503***	-0.448***	-0.595***	-0.569***
	(0.116)	(0.128)	(0.133)	(0.160)	(0.158)	(0.161)	(0.170)	(0.174)	(0.182)
RTC-Unskilled$_r$ $\times$ Dist. L.O.$_r$	0.163	0.176	0.237**	0.252**	0.326***	0.251**	0.248**	0.290***	0.330***
	(0.111)	(0.110)	(0.118)	(0.119)	(0.106)	(0.098)	(0.098)	(0.095)	(0.092)
Share female$_{1991}$		0.401	0.933**	0.946**	0.611	0.277	0.249	0.119	-0.148
		(0.317)	(0.388)	(0.385)	(0.403)	(0.365)	(0.365)	(0.369)	(0.447)
Share high-skill$_{1991}$			-0.293***	-0.222**	-0.311***	-0.276***	-0.310***	-0.426***	-0.442***
			(0.092)	(0.091)	(0.084)	(0.076)	(0.084)	(0.089)	(0.085)
Share urban$_{1991}$				-0.089***	-0.083***	-0.075**	-0.075**	-0.022	-0.026
				(0.034)	(0.032)	(0.029)	(0.030)	(0.032)	(0.033)
log($population_{1991}$)					0.016***	0.013***	0.013***	0.010***	0.011***
					(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Informality$_{1980}$						0.248***	0.245***	0.258***	0.259***
						(0.028)	(0.028)	(0.028)	(0.029)
$Gini_{1991}$							0.046	0.017	0.018
							(0.047)	(0.047)	(0.048)
log (Govt. Spending$_{1991}$)								-0.024***	-0.026***
								(0.007)	(0.007)
Distance to Capital									-0.001
									(0.001)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.380	0.383	0.398	0.407	0.434	0.534	0.535	0.545	0.564

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.5: Effects on Informality: High-Skill workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC-Skilled$_r$	-0.147	-0.123	-0.034	0.021	-0.205	-0.336	-0.135	-0.251	-0.040
	(0.289)	(0.322)	(0.321)	(0.334)	(0.323)	(0.326)	(0.341)	(0.341)	(0.338)
RTC-Skilled$_r$ $\times$ Dist. L.O.$_r$	0.124	0.118	0.089	0.073	0.148	0.190	0.120	0.203	0.019
	(0.197)	(0.201)	(0.199)	(0.207)	(0.199)	(0.202)	(0.197)	(0.197)	(0.181)
Share female$_{1991}$		-0.090	0.231	0.240	-0.127	-0.194	-0.386	-0.606	-0.848
		(0.335)	(0.425)	(0.428)	(0.447)	(0.454)	(0.457)	(0.466)	(0.549)
Share high-skill$_{1991}$			-0.105	-0.083	-0.173*	-0.184*	-0.295***	-0.483***	-0.453***
			(0.089)	(0.096)	(0.092)	(0.097)	(0.101)	(0.118)	(0.116)
Share urban$_{1991}$				-0.019	-0.023	-0.021	-0.007	0.060	0.050
				(0.042)	(0.039)	(0.040)	(0.041)	(0.044)	(0.044)
$\log(population_{1991})$					0.015***	0.015***	0.012***	0.008*	0.007
					(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Informality$_{1980}$						0.220	0.202	0.250	0.267*
						(0.164)	(0.153)	(0.153)	(0.161)
$Gini_{1991}$							0.215***	0.182**	0.192***
							(0.072)	(0.071)	(0.072)
$\log$ (Govt. Spending$_{1991}$)								-0.034***	-0.035***
								(0.010)	(0.011)
Distance to Capital									-0.000
									(0.002)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.306	0.306	0.308	0.309	0.327	0.334	0.352	0.366	0.346

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.6: Effects on Non-employment: All workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC_r	0.426***	0.506***	0.464***	0.382***	0.348***	0.354***	0.375***	0.366***	0.349***
	(0.066)	(0.072)	(0.080)	(0.100)	(0.106)	(0.104)	(0.107)	(0.106)	(0.116)
$\text{RTC}_r \times \text{Dist. L.O.}_r$	-0.071	-0.082	-0.090	-0.091	-0.083	-0.084	-0.086	-0.083	-0.059
	(0.054)	(0.056)	(0.055)	(0.064)	(0.066)	(0.066)	(0.067)	(0.066)	(0.068)
Share female_{1991}		-0.350**	-0.424***	-0.433***	-0.489***	-0.519***	-0.526***	-0.538***	-0.565***
		(0.142)	(0.159)	(0.159)	(0.165)	(0.169)	(0.167)	(0.166)	(0.198)
Share high-skill_{1991}			0.048	0.022	0.010	0.010	-0.004	-0.015	-0.042
			(0.046)	(0.050)	(0.052)	(0.051)	(0.058)	(0.060)	(0.061)
Share urban_{1991}				0.032	0.032	0.033	0.033	0.038*	0.043*
				(0.020)	(0.020)	(0.021)	(0.021)	(0.022)	(0.023)
$\log(population_{1991})$					0.003	0.003	0.003	0.003	0.003
					(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\text{Non-employment}_{1980}$						0.029	0.026	0.028	0.024
						(0.042)	(0.042)	(0.042)	(0.042)
$Gini_{1991}$							0.018	0.016	0.031
							(0.026)	(0.026)	(0.027)
$\log(\text{Govt. Spending}_{1991})$								-0.002	-0.001
								(0.005)	(0.004)
Distance to Capital									-0.000
									(0.001)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.396	0.407	0.409	0.415	0.418	0.419	0.420	0.416	0.416

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.7: Effects on Non-employment: Low-Skill workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC-Unskilled$_r$	0.502***	0.598***	0.578***	0.496***	0.453***	0.460***	0.480***	0.469***	0.459***
	(0.070)	(0.077)	(0.087)	(0.107)	(0.113)	(0.112)	(0.115)	(0.115)	(0.126)
RTC-Unskilled$_r$ $\times$ Dist. L.O.$_r$	-0.132**	-0.147**	-0.151**	-0.159**	-0.148**	-0.149**	-0.151**	-0.148**	-0.122
	(0.060)	(0.061)	(0.062)	(0.071)	(0.073)	(0.073)	(0.074)	(0.073)	(0.077)
Share female$_{1991}$		-0.420***	-0.455***	-0.464***	-0.533***	-0.567***	-0.575***	-0.587***	-0.602***
		(0.151)	(0.173)	(0.173)	(0.179)	(0.184)	(0.181)	(0.180)	(0.218)
Share high-skill$_{1991}$			0.023	-0.008	-0.024	-0.030	-0.043	-0.055	-0.081
			(0.051)	(0.055)	(0.056)	(0.056)	(0.063)	(0.065)	(0.067)
Share urban$_{1991}$				0.036*	0.036	0.037*	0.037*	0.042*	0.046*
				(0.021)	(0.022)	(0.022)	(0.022)	(0.023)	(0.025)
$\log(population_{1991})$					0.003	0.004*	0.003	0.003	0.004
					(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Non-employment$_{1980}$						0.032	0.030	0.032	0.029
						(0.043)	(0.043)	(0.043)	(0.044)
$Gini_{1991}$							0.018	0.016	0.028
							(0.028)	(0.028)	(0.029)
$\log$ (Govt. Spending$_{1991}$)								-0.002	-0.002
								(0.005)	(0.005)
Distance to Capital									-0.000
									(0.001)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.405	0.418	0.418	0.424	0.428	0.429	0.430	0.428	0.426

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.8: Effects on Non-employment: High-Skill workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC-Skilled$_r$	0.132	0.139	0.172	0.140	0.121	0.122	0.175	0.173	0.025
	(0.098)	(0.100)	(0.105)	(0.124)	(0.127)	(0.128)	(0.130)	(0.129)	(0.102)
RTC-Skilled$_r$ $\times$ Dist. L.O.$_r$	0.001	-0.001	-0.011	-0.002	0.004	0.016	-0.002	-0.001	0.079
	(0.081)	(0.081)	(0.082)	(0.088)	(0.089)	(0.090)	(0.092)	(0.091)	(0.054)
Share female$_{1991}$		-0.025	0.078	0.076	0.043	0.079	0.032	0.028	-0.092
		(0.127)	(0.157)	(0.157)	(0.159)	(0.163)	(0.171)	(0.173)	(0.183)
Share high-skill$_{1991}$			-0.037	-0.052	-0.060	-0.052	-0.085	-0.089	-0.112**
			(0.039)	(0.045)	(0.047)	(0.047)	(0.052)	(0.061)	(0.054)
Share urban$_{1991}$				0.012	0.011	0.016	0.019	0.021	0.041**
				(0.016)	(0.016)	(0.017)	(0.017)	(0.020)	(0.018)
log($population_{1991}$)					0.001	0.001	0.000	-0.000	-0.000
					(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Non-employment$_{1980}$						-0.097*	-0.100*	-0.099*	-0.141**
						(0.055)	(0.056)	(0.056)	(0.059)
$Gini_{1991}$							0.061**	0.061**	0.093***
							(0.026)	(0.026)	(0.025)
log (Govt. Spending$_{1991}$)								-0.001	0.000
								(0.005)	(0.005)
Distance to Capital									-0.003***
									(0.001)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.325	0.325	0.327	0.327	0.328	0.336	0.346	0.343	0.379

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.9: Effects on Wages: All workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC_r	-0.384*	-0.443**	-0.954***	-1.052***	-0.846***	-0.793***	-0.148	-0.048	-0.187
	(0.200)	(0.224)	(0.247)	(0.263)	(0.281)	(0.278)	(0.275)	(0.277)	(0.297)
$\text{RTC}_r \times$ Dist. L.O.$_r$	0.091	0.098	-0.016	-0.016	-0.078	-0.094	-0.166	-0.205	-0.275*
	(0.164)	(0.164)	(0.189)	(0.178)	(0.184)	(0.185)	(0.157)	(0.155)	(0.161)
Share female$_{1991}$		0.258	-0.812*	-0.820*	-0.473	-0.782*	-1.124**	-0.977**	-1.040*
		(0.419)	(0.450)	(0.452)	(0.467)	(0.467)	(0.500)	(0.495)	(0.543)
Share high-skill$_{1991}$			0.624***	0.595***	0.669***	0.671***	0.248	0.375**	0.437**
			(0.142)	(0.143)	(0.138)	(0.137)	(0.154)	(0.168)	(0.172)
Share urban$_{1991}$				0.038	0.037	0.051	0.041	-0.013	0.002
				(0.053)	(0.055)	(0.055)	(0.047)	(0.054)	(0.057)
$\log(population_{1991})$					-0.016***	-0.012**	-0.019***	-0.017***	-0.019***
					(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Non-employment$_{1980}$						0.315**	0.246**	0.224**	0.206*
						(0.125)	(0.113)	(0.113)	(0.114)
$Gini_{1991}$							0.573***	0.596***	0.628***
							(0.082)	(0.082)	(0.085)
log (Govt. Spending$_{1991}$)								0.025*	0.028**
								(0.013)	(0.014)
Distance to Capital									-0.005**
									(0.002)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.583	0.583	0.603	0.604	0.612	0.619	0.671	0.672	0.667

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.10: Effects on Wages: Low-Skill workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC-Unskilled$_r$	-0.218	-0.144	-0.641**	-0.928***	-0.710**	-0.654**	-0.054	0.111	0.003
	(0.216)	(0.241)	(0.275)	(0.278)	(0.297)	(0.293)	(0.290)	(0.291)	(0.314)
RTC-Unskilled$_r$ $\times$ Dist. L.O.$_r$	0.058	0.049	-0.082	-0.099	-0.169	-0.185	-0.237	-0.298*	-0.373**
	(0.204)	(0.200)	(0.241)	(0.199)	(0.202)	(0.201)	(0.176)	(0.171)	(0.175)
Share female$_{1991}$		-0.322	-1.417***	-1.438***	-1.086**	-1.445***	-1.764***	-1.551***	-1.702***
		(0.452)	(0.489)	(0.488)	(0.506)	(0.509)	(0.567)	(0.556)	(0.599)
Share high-skill$_{1991}$			0.632***	0.532***	0.615***	0.568***	0.169	0.363**	0.422**
			(0.157)	(0.154)	(0.145)	(0.142)	(0.161)	(0.171)	(0.174)
Share urban$_{1991}$				0.122**	0.120**	0.137**	0.129**	0.046	0.061
				(0.056)	(0.058)	(0.059)	(0.052)	(0.058)	(0.062)
$\log(population_{1991})$					-0.016***	-0.011*	-0.019***	-0.015***	-0.018***
					(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Non-employment$_{1980}$						0.357***	0.288**	0.259**	0.239**
						(0.130)	(0.120)	(0.119)	(0.121)
$Gini_{1991}$							0.556***	0.592***	0.630***
							(0.089)	(0.089)	(0.093)
$\log(\text{Govt. Spending}_{1991})$								0.038***	0.040***
								(0.014)	(0.014)
Distance to Capital									-0.005*
									(0.002)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.525	0.526	0.547	0.553	0.562	0.571	0.620	0.628	0.620

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.11: Effects on Wages: High-Skill workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC-Skilled$_r$	-1.410***	-1.363***	-1.224***	-0.576	-0.066	-0.070	-0.119	-0.137	-0.184
	(0.360)	(0.387)	(0.415)	(0.445)	(0.453)	(0.445)	(0.481)	(0.483)	(0.527)
RTC-Skilled$_r$ $\times$ Dist. L.O.$_r$	0.277	0.266	0.223	0.034	-0.139	-0.207	-0.191	-0.177	-0.201
	(0.272)	(0.274)	(0.270)	(0.298)	(0.299)	(0.295)	(0.309)	(0.315)	(0.349)
Share female$_{1991}$		-0.173	0.309	0.388	1.222*	0.997	1.045	1.007	1.456*
		(0.513)	(0.704)	(0.687)	(0.680)	(0.702)	(0.715)	(0.734)	(0.832)
Share high-skill$_{1991}$			-0.166	0.125	0.325	0.285	0.313	0.277	0.286
			(0.196)	(0.256)	(0.299)	(0.292)	(0.315)	(0.332)	(0.341)
Share urban$_{1991}$				-0.242***	-0.234***	-0.259***	-0.262***	-0.249***	-0.263***
				(0.078)	(0.084)	(0.083)	(0.086)	(0.091)	(0.096)
log($population_{1991}$)					-0.033***	-0.030***	-0.029***	-0.030***	-0.031***
					(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Non-employment$_{1980}$						0.556**	0.558**	0.561**	0.613**
						(0.236)	(0.235)	(0.239)	(0.255)
$Gini_{1991}$							-0.055	-0.060	-0.053
							(0.109)	(0.109)	(0.116)
log (Govt. Spending$_{1991}$)								-0.007	-0.002
								(0.019)	(0.019)
Distance to Capital									-0.001
									(0.004)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.556	0.556	0.557	0.572	0.596	0.605	0.605	0.599	0.603

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.12: Effects on Self-Employment: All workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC_r	0.512***	0.618***	0.414***	-0.050	-0.070	-0.142	-0.148	-0.096	-0.179
	(0.084)	(0.096)	(0.118)	(0.126)	(0.137)	(0.141)	(0.140)	(0.142)	(0.152)
$\text{RTC}_r \times \text{Dist. L.O.}_r$	0.432***	0.420***	0.376***	0.385***	0.390***	0.387***	0.387***	0.373***	0.362***
	(0.101)	(0.093)	(0.109)	(0.077)	(0.080)	(0.080)	(0.080)	(0.081)	(0.080)
Share female_{1991}		-0.460*	-0.873***	-0.910***	-0.943***	-0.991***	-0.988***	-0.936***	-1.187***
		(0.239)	(0.292)	(0.273)	(0.292)	(0.288)	(0.292)	(0.287)	(0.290)
Share high-skill_{1991}			0.243***	0.112*	0.105*	0.104*	0.107	0.155**	0.153**
			(0.070)	(0.061)	(0.063)	(0.063)	(0.069)	(0.071)	(0.073)
Share urban_{1991}				0.179***	0.179***	0.180***	0.180***	0.160***	0.185***
				(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.030)
$\log(population_{1991})$					0.001	0.001	0.001	0.002	0.002
					(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$\text{Informality}_{1980}$						0.030	0.030	0.025	0.009
						(0.028)	(0.028)	(0.029)	(0.030)
$Gini_{1991}$							-0.004	0.005	0.059
							(0.042)	(0.043)	(0.043)
$\log(\text{Govt. Spending}_{1991})$								0.009	0.012*
								(0.006)	(0.007)
Distance to Capital									-0.005***
									(0.001)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.640	0.644	0.655	0.698	0.698	0.700	0.700	0.701	0.725

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.13: Effects on Self-Employment: Low-Skill workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC-Unskilled$_r$	0.673***	0.746***	0.460***	0.001	-0.029	-0.212	-0.228	-0.158	-0.219
	(0.097)	(0.112)	(0.135)	(0.139)	(0.151)	(0.154)	(0.153)	(0.155)	(0.165)
RTC-Unskilled$_r$ × Dist. L.O.$_r$	0.533***	0.523***	0.455***	0.416***	0.425***	0.409***	0.410***	0.393***	0.377***
	(0.114)	(0.109)	(0.131)	(0.093)	(0.096)	(0.097)	(0.096)	(0.097)	(0.098)
Share female$_{1991}$		-0.322	-0.912***	-0.953***	-1.001***	-1.096***	-1.088***	-1.024***	-1.353***
		(0.270)	(0.336)	(0.313)	(0.333)	(0.334)	(0.338)	(0.331)	(0.329)
Share high-skill$_{1991}$			0.351***	0.187**	0.175**	0.183**	0.194**	0.253***	0.246***
			(0.084)	(0.073)	(0.076)	(0.075)	(0.083)	(0.083)	(0.086)
Share urban$_{1991}$				0.198***	0.198***	0.201***	0.201***	0.174***	0.200***
				(0.031)	(0.032)	(0.032)	(0.032)	(0.032)	(0.034)
log($population_{1991}$)					0.002	0.001	0.002	0.003	0.003
					(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Informality$_{1980}$						0.069**	0.070**	0.064**	0.048
						(0.027)	(0.027)	(0.029)	(0.031)
$Gini_{1991}$							-0.013	-0.002	0.057
							(0.046)	(0.047)	(0.048)
log (Govt. Spending$_{1991}$)								0.012*	0.015*
								(0.007)	(0.008)
Distance to Capital									-0.005***
									(0.002)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.692	0.694	0.708	0.744	0.744	0.750	0.750	0.752	0.769

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.14: Effects on Self-Employment: High-Skill workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTC-Skilled$_r$	0.220***	0.199**	0.102	0.137	0.143	0.112	0.131	0.135	0.100
	(0.072)	(0.078)	(0.078)	(0.089)	(0.092)	(0.093)	(0.093)	(0.094)	(0.096)
RTC-Skilled$_r$ × Dist. L.O.$_r$	-0.053	-0.048	-0.020	-0.030	-0.032	-0.022	-0.029	-0.032	-0.033
	(0.051)	(0.052)	(0.050)	(0.052)	(0.052)	(0.052)	(0.051)	(0.053)	(0.055)
Share female$_{1991}$		0.074	-0.248*	-0.246*	-0.236	-0.252*	-0.270*	-0.261*	-0.368**
		(0.103)	(0.139)	(0.140)	(0.144)	(0.148)	(0.152)	(0.153)	(0.152)
Share high-skill$_{1991}$			0.112***	0.128***	0.130***	0.126***	0.116***	0.124***	0.117***
			(0.026)	(0.027)	(0.029)	(0.030)	(0.032)	(0.036)	(0.035)
Share urban$_{1991}$				-0.013	-0.013	-0.012	-0.011	-0.014	-0.008
				(0.012)	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)
log($population_{1991}$)					-0.000	-0.000	-0.001	-0.000	-0.000
					(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Informality$_{1980}$						0.060*	0.058*	0.057	0.050
						(0.034)	(0.034)	(0.035)	(0.037)
$Gini_{1991}$							0.021	0.022	0.048*
							(0.024)	(0.025)	(0.025)
log (Govt. Spending$_{1991}$)								0.001	0.002
								(0.003)	(0.003)
Distance to Capital									-0.002***
									(0.001)
Observations	411	411	411	411	411	411	411	410	395
R-squared	0.157	0.158	0.185	0.186	0.187	0.192	0.194	0.191	0.224

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects.

Table A.15: Effects on Informal Employees, All Employees and Self-Employed as Share of Working Age Population

Sample (by workers' skill level):	All			Low			High		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Informal Employees	All Employees	Self Employed	Informal Employees	All Employees	Self Employed	Informal Employees	All Employees	Self Employed
RTC_r	-0.011	-0.078	-0.131						
	(0.121)	(0.213)	(0.113)						
RTC × Dist. labor office	0.174***	0.331***	0.349***						
	(0.056)	(0.099)	(0.052)						
RTC-Unskilled				-0.036	-0.401**	-0.194*			
				(0.110)	(0.188)	(0.112)			
RTC-Unskilled × Dist. labor office				0.206***	0.313***	0.386***			
				(0.054)	(0.093)	(0.056)			
RTC-Skilled							0.050	0.615	0.239
							(0.433)	(0.628)	(0.155)
RTC-Skilled × Dist. labor office							-0.990***	-0.749**	-0.060
							(0.236)	(0.343)	(0.085)
Observations	395	395	395	395	395	395	395	395	395
R-squared	0.580	0.626	0.635	0.568	0.741	0.576	0.806	0.772	0.457

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects. The controls used are the following: share of women, high-skill individuals, urban population and log-population in 1991; informality and unemployment levels in 1980; logarithm of local government spending (per capita) and the Gini coefficient, both measured in 1991; and median driving distance to the state's capital.

Table A.16: Effects on Informality and Wages at the Industry-by-MMC level

	Informality			Wages		
Sample (by workers' skill level):	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta Tariffs_s \times$ Dist. L.O.$_r$	0.090***	0.102***	0.073***	-0.183***	-0.215***	-0.107**
	(0.013)	(0.014)	(0.025)	(0.018)	(0.018)	(0.043)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,858	5,621	3,829	6,341	6,102	4,469
R-squared	0.601	0.584	0.680	0.642	0.661	0.579

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions control for industry and micro-region fixed effects, and the following demographic controls: share of women, high-skill individuals, urban population and log-population in 1991.

Table A.17: Effects on Informality and Non-Employment using Mean Distance to L.O.

	Informality			Non-Employment		
Sample (by workers' skill level):	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
RTC_r	0.341**			0.316***		
	(0.147)			(0.100)		
$RTC_r \times$ Distance L.O.$_r$	0.184*			-0.084		
	(0.109)			(0.082)		
RTC-Unskilled$_r$		0.326**			0.407***	
		(0.149)			(0.107)	
RTC-Unsk.$_r$ $\times$ Distance L.O.$_r$		0.309**			-0.157*	
		(0.123)			(0.093)	
RTC-Skilled$_r$			-0.236			0.162
			(0.293)			(0.126)
RTC-Skill.$_r$ $\times$ Distance L.O.$_r$			0.237			-0.029
			(0.245)			(0.116)
Observations	411	411	411	411	411	411
R-squared	0.393	0.429	0.331	0.409	0.419	0.326

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow the specification in expression 5, which includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and the state fixed effects. The demographic controls are: share of women, high-skill individuals, urban population and log-population in 1991.

Table A.18: Effects on Informality, Non-Employment and Wages using Imports as Trade Shock

	Informality			Non-employment			Wages		
Sample (by workers' skill level):	All	Low	High	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Imports	-0.126	0.013	0.016	0.542**	0.911***	0.261	-2.670**	-2.463**	-1.370
	(0.421)	(0.457)	(0.608)	(0.258)	(0.273)	(0.245)	(1.049)	(1.063)	(1.089)
Imports $\times$ Dist. L.O.$_r$	1.488***	1.697***	-0.400	-0.135	-0.441	-0.015	0.340	0.155	-1.167
	(0.462)	(0.512)	(0.614)	(0.255)	(0.272)	(0.236)	(0.842)	(0.889)	(1.195)
Observations	411	411	411	411	411	411	411	411	411
R-squared	0.405	0.447	0.327	0.404	0.411	0.326	0.618	0.567	0.606

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow a specification analogous to expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and state fixed effects. The demographic controls are: share of women, high-skill individuals, urban population and log-population in 1991.

Table A.19: Effects on Informality, Non-Employment and Wages using Penetration Coefficient as Trade Shock

	Informality			Non-employment			Wages		
Sample (by workers' skill level):	All	Low	High	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Penetration Coef.	0.234	0.483	-0.052	0.489*	0.856***	0.198	-3.330***	-3.170***	-2.111*
	(0.481)	(0.519)	(0.681)	(0.284)	(0.308)	(0.267)	(1.130)	(1.156)	(1.211)
Penet. Coef. × Dist. L.O.$_r$	1.153**	1.279**	-0.472	-0.139	-0.446	0.015	0.470	0.314	-0.803
	(0.460)	(0.506)	(0.621)	(0.252)	(0.271)	(0.244)	(0.813)	(0.868)	(1.217)
Observations	411	411	411	411	411	411	411	411	411
R-squared	0.401	0.442	0.328	0.401	0.406	0.325	0.621	0.571	0.608

Notes: Robust standard errors reported. Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. All regressions follow a specification analogous to expression 4, which also includes the interaction between the number of inspectors at the state level ($Inspectors_s$) and distance to the labor office, $Dist_r$, the distance in levels ($Dist_r$) and state fixed effects. The demographic controls are: share of women, high-skill individuals, urban population and log-population in 1991.

D Additional results

Table A.20: IV Estimation – First stage results

Sample:	All Workers		Low-Skill		High-Skill	
Endogenous regressor:	$Enforce_r$	$RTC_r \times Enforce_r$	$Enforce_r$	RTC_r-Unsk.$\times Enforce_r$	$Enforce_r$	RTC_r-Skill.$\times Enforce_r$
	(1)	(2)	(3)	(4)	(5)	(6)
RTC_r	-48.984***	6.580***				
	(16.000)	(1.419)				
$RTC_r \times$ Dist. L.O.	54.501***	1.563				
	(11.496)	(1.216)				
Distance L.O. (per 100km)	-3.106***	-0.193***	-3.067***	-0.183***	-3.853***	-0.451***
	(0.687)	(0.057)	(0.686)	(0.058)	(1.100)	(0.126)
$Inspectors_s \times$ Distance L.O.	0.007*	0.002***	0.007*	0.002***	0.010***	0.002***
	(0.004)	(0.000)	(0.004)	(0.000)	(0.004)	(0.000)
RTC_r-Unskilled			-53.725***	6.798***		
			(15.692)	(1.457)		
RTC_r-Unsk.$\times$ Dist. L.O.			56.087***	1.287		
			(12.324)	(1.319)		
RTC_r-Skilled					-2.746	6.875**
					(21.792)	(2.703)
RTC_r-Skill.$\times$ Dist. L.O.					35.656***	3.263**
					(12.925)	(1.508)
Observations	411	411	411	411	411	411
R-squared	0.633	0.746	0.621	0.738	0.687	0.705

Notes: Significant at the *** 1 percent, ** 5 percent, and * 10 percent level. Estimates obtained using the limited information maximum likelihood estimator. All regressions control for state fixed effects and the following demographic controls: share of women, high-skill individuals, urban population and log-population in 1991.