

# DISCUSSION PAPER SERIES

DP18860

## **TRADE AND DOMESTIC DISTORTIONS: THE CASE OF INFORMALITY**

Rafael Dix-Carneiro, Pinelopi Goldberg, Costas  
Meghir and Gabriel Ulyssea

**INTERNATIONAL TRADE AND  
REGIONAL ECONOMICS AND LABOUR  
ECONOMICS**



# TRADE AND DOMESTIC DISTORTIONS: THE CASE OF INFORMALITY

*Rafael Dix-Carneiro, Pinelopi Goldberg, Costas Meghir and Gabriel Ulyssea*

Discussion Paper DP18860  
Published 24 February 2024  
Submitted 21 February 2024

Centre for Economic Policy Research  
33 Great Sutton Street, London EC1V 0DX, UK  
Tel: +44 (0)20 7183 8801  
[www.cepr.org](http://www.cepr.org)

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- International Trade and Regional Economics
- Labour Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Rafael Dix-Carneiro, Pinelopi Goldberg, Costas Meghir and Gabriel Ulyssea

# TRADE AND DOMESTIC DISTORTIONS: THE CASE OF INFORMALITY

## Abstract

We examine the effects of international trade in the presence of a set of domestic distortions giving rise to informality, a prevalent phenomenon in developing countries. In our quantitative model, the informal sector arises from burdensome taxes and regulations that are imperfectly enforced by the government. Consequently, smaller, less productive firms face fewer distortions than larger, more productive ones, potentially leading to substantial misallocation. We show that in settings with a large informal sector, the gains from trade are significantly amplified, as reductions in trade barriers imply a reallocation of resources from initially less distorted to more distorted firms. We confirm findings from earlier reduced-form studies that the informal sector mitigates the impact of negative labor demand shocks on unemployment. Nonetheless, the informal sector can exacerbate the adverse welfare effects of economic downturns, amplifying misallocation. Last, our research sheds light on the relationship between trade openness and cross-firm wage inequality.

JEL Classification: F14, F16, J46, O17

Keywords: Informality

Rafael Dix-Carneiro - rafael.dix.carneiro@duke.edu  
*Department of Economics, Duke University*

Pinelopi Goldberg - penny.goldberg@yale.edu  
*Yale University and NBER and CEPR*

Costas Meghir - c.meghir@yale.edu  
*Yale University and NBER and CEPR*

Gabriel Ulyssea - g.ulyssea@ucl.ac.uk  
*University College London and CEPR*

## Acknowledgements

This is a substantially revised version of "Trade and Informality in the Presence of Labor Market Frictions and Regulations." This project was supported by award SES-1629124 from the National Science Foundation and by the Early Career Research Grant 15-150-04 from the W.E. Upjohn Institute for Employment Research. We would like to thank Federico Huneus, Nina Pavcnik, Jean-Marc Robin, Sharon Traiberman, and Jim Tybout, as well as numerous seminar and conference participants, for helpful comments and discussions.

# TRADE AND DOMESTIC DISTORTIONS: THE CASE OF INFORMALITY\*

Rafael Dix-Carneiro, Duke University  
Pinelopi Koujianou Goldberg, Yale University  
Costas Meghir, Yale University  
Gabriel Ulyssea, University College London

February 13, 2024

## Abstract

We examine the effects of international trade in the presence of a set of domestic distortions giving rise to informality, a prevalent phenomenon in developing countries. In our quantitative model, the informal sector arises from burdensome taxes and regulations that are imperfectly enforced by the government. Consequently, smaller, less productive firms face fewer distortions than larger, more productive ones, potentially leading to substantial misallocation. We show that in settings with a large informal sector, the gains from trade are significantly amplified, as reductions in trade barriers imply a reallocation of resources from initially less distorted to more distorted firms. We confirm findings from earlier reduced-form studies that the informal sector mitigates the impact of negative labor demand shocks on unemployment. Nonetheless, the informal sector can exacerbate the adverse welfare effects of economic downturns, amplifying misallocation. Last, our research sheds light on the relationship between trade openness and cross-firm wage inequality.

*JEL* codes: F14, F16, J46, O17

## 1 Introduction

In the past 40 years, most developing countries have opened up to foreign competition, which has long been seen as a necessary step to modernize their economies and promote economic growth. However, these economies are also characterized by the presence of severe domestic frictions (e.g., [Banerjee and Dufo, 2005](#); [Hsieh and Klenow, 2009](#)), which can shape the effects of international trade in important ways by either amplifying or mitigating the impacts of trade liberalization ([Atkin and Khandelwal, 2020](#); [Atkin and Donaldson, 2022](#); [Bai et al., 2023](#)). Thus, to understand

---

\*Dix-Carneiro: rafael.dix.carneiro@duke.edu; Goldberg: penny.goldberg@yale.edu; Meghir: c.meghir@yale.edu; Ulyssea: g.ulyssea@ucl.ac.uk. This is a substantially revised version of “Trade and Informality in the Presence of Labor Market Frictions and Regulations.” This project was supported by award SES-1629124 from the National Science Foundation and by the Early Career Research Grant 15-150-04 from the W.E. Upjohn Institute for Employment Research. We would like to thank Federico Huneeus, Nina Pavcnik, Jean-Marc Robin, Sharon Traiberman, and Jim Tybout, as well as numerous seminar and conference participants, for helpful comments and discussions.

whether trade hinders or promotes economic development, it is key to understand how it interacts with domestic frictions in developing countries.

This paper focuses on a set of domestic distortions that give rise to a phenomenon that is pervasive in developing economies: informality. The informal sector is broadly defined as the part of the economy (firms and/or workers) that evades burdensome taxes and labor market regulations. Remarkably, a substantial share of the labor force in many emerging and developing countries is employed informally; for example, in Latin America, the informally employed share of the labor force ranges from 35% in Chile to 80% in Peru (Perry et al., 2007). Given the informal sector’s sheer magnitude and central role in the functioning of these economies, it is a priori likely that informality plays an important role in shaping the aggregate and distributional effects of trade. Indeed, recent empirical work has emphasized that shifts into or out of informality represent significant margins through which economies adjust to changes in the global environment.<sup>1</sup> However, the informal sector has received little attention in theoretical and empirical studies of international trade.

To model the impacts of trade in an economy with a large informal sector, we start with the premise that it arises from burdensome taxes and labor market regulations imposed on firms, which the government can only imperfectly enforce. Although the government imposes costs on informal firms through fines or the possibility of closure, these measures are not consistently applied to all firms. This lack of strict enforcement creates incentives for certain firms to operate informally, evading taxes, and bypassing compliance with labor market regulations. Importantly, this leads to a situation where firms face unequal exposure to distorting taxes and regulations.

Our framework considers a rich institutional setting, in which the government imposes onerous labor market regulations that discourage formal status. Past research has highlighted the importance of regulations such as firing costs and minimum wages in driving the size of the informal sector in Latin American countries (e.g., Heckman and Pagés, 2000). Moreover, formal firms that are legally registered with the authorities face value-added and payroll taxes, which further reduce their profitability. Finally, imported goods are subject to import tariffs.

The selection of the remaining ingredients in our model is guided by a number of facts on formal and informal firms and workers in Brazil.<sup>2</sup> In particular, the proportion of firms that are informal falls substantially with firm size, indicating that informal firms perceive increasing costs associated with informality as they grow. For example, as informal firms grow, they become more visible to the local authorities, leading to increased fines and threat of closure. Concurrently, in order to grow, informal firms face larger costs of capital (Catão et al., 2009), indicating larger opportunity costs of informality for larger firms. In addition, informal sector firms tend to be substantially smaller and exhibit lower worker productivity than their formal sector counterparts. These facts motivate a model based on firms that are heterogeneous in their total factor productivity and sort into the formal and informal sectors accordingly.

---

<sup>1</sup>For example, see Goldberg and Pavcnik (2003), McCaig and Pavcnik (2018), Dix-Carneiro and Kovak (2019), Ponczek and Ulyssea (2022).

<sup>2</sup>See Meghir et al. (2015) and Ulyssea (2020) for a discussion of a large set of stylized facts involving the informal sector in Brazil.

If a firm is productive and large enough, it faces prohibitive costs of operating informally and sorts into the formal sector. However, less productive and smaller firms tend to sort into the informal sector, flying under the government’s radar and avoiding the burdens of distortive taxes and regulations. In simple terms, this environment creates an economy with a size-dependent distortion: the larger (and more productive) firms in the formal sector are subject to taxes and regulations and experience larger distortions compared to the smaller scale (and lower productivity) informal sector. This implies that formal firms underproduce relative to the social optimum whereas informal firms overproduce.

Search and matching frictions in the labor market generate equilibrium unemployment and wage dispersion. These elements are important for three reasons. First, in the data, transitions from unemployment to informality are twice as likely as transitions from unemployment to formality. This implies that any policy or shock that leads to higher inflows into unemployment is also likely to lead to higher informality. Second, data on formal and informal firms reveal that the latter tend to pay lower wages—often below the official minimum wage—than their formal counterparts. This equilibrium wage dispersion is crucial for ensuring that the minimum wage meaningfully impacts firms’ incentives to opt for either the formal or informal sector. Finally, we are motivated by the work by [Dix-Carneiro and Kovak \(2019\)](#) and [Ponczek and Ulyssea \(2022\)](#) who show the important role of the informal sector in mitigating the unemployment consequences of trade-induced negative labor demand shocks.

Our model encompasses two broad sectors: manufacturing and services. It is a priori important to carefully model both sectors for two reasons. The first is that informality is widespread in services. Second, although we do not allow for trade in service sector goods, service sector firms are also affected by trade through intersectoral linkages with manufacturing. This creates a potentially important channel linking trade to the aggregate level of informality. In turn, international trade plays a role in the model through two channels. First, imports impact all firms in the economy because of their direct impact on the price of imported manufacturing goods and the aforementioned intersectoral linkages. Second, manufacturing formal firms can export, subject to fixed export costs and variable trade costs (as in [Melitz, 2003](#)).

We estimate the model using multiple data sources, including matched employer-employee data from formal and informal firms and workers in Brazil and several other sources of firm- and worker-level data such as household surveys, manufacturing and services censuses, and customs data. Brazil provides an excellent setting for our work. Not only does it offer outstanding data sources for both the formal and informal sectors, but it also presents a particularly relevant environment given its large informal sector. Nearly two-thirds of firms and 50% of workers are informal, and the informal sector generates 40% of GDP ([Ulyssea, 2018](#)). Moreover, there is a clear definition of what constitutes informality in the country: we define as informal workers those who do not hold a formal labor contract, clearly observable through the *carteira de trabalho*, a booklet logging each worker’s (formal) employment history. Informal firms are those not registered with the tax authorities, indicated by the absence of a tax identification number (*Cadastro Nacional de Pessoa*

*Jurídica*—CNPJ), which we can observe.

Armed with our model estimates, we start by analyzing the effects of a reduction in bilateral trade barriers on the Brazilian economy. Considering the event of a small globalization shock, we are able to decompose the impact on real income into a “mechanical effect” and a “reallocation effect” (similar to [Baqae and Farhi, 2020](#); [Atkin and Donaldson, 2022](#)). The mechanical effect measures the real income benefit arising from a reduction in the prices paid by consumers and firms for imported goods, where labor allocations, wages, aggregate income, and the demand that all firms face are held constant. It isolates the immediate effects on real income resulting from lower trade costs in the absence of any reallocation. On the other hand, the reallocation effect reflects the effects of reallocating resources across uses that are more or less distorted—in this case with trade costs and technology held constant—along with other general equilibrium effects.<sup>3</sup> As is well known, the reallocation effect would be zero in an economy with no distortions ([Baqae and Farhi, 2020](#)). Strikingly, though, we find that the reallocation effect is almost five times as large as the mechanical effect, suggesting that the positive effects of trade liberalization operate mainly through the reallocation of resources to more efficient uses and other general equilibrium effects. Furthermore, these reallocation effects are positive and large even though the unemployment rate increases.

To better understand these results, we examine how informality, measured aggregate total factor productivity (TFP), and the dispersion of the log marginal revenue product of labor respond to changes in trade costs.<sup>4</sup> A reduction in trade costs leads to a contraction of informality in both the manufacturing and service sectors. This contraction of the less productive and less distorted informal sector is accompanied by a reallocation of resources to formal firms which are initially more productive and more distorted, leading to an increase in measured aggregate TFP. Simultaneously, we observe a decline in the dispersion of the log marginal revenue product of labor, indicating a more efficient allocation of resources following increases in globalization, further boosting the reallocation effect ([Hsieh and Klenow, 2009](#)). All these results suggest that reductions in trade costs trigger a reallocation of resources toward more efficient uses and that this reallocation has an important quantitative impact on real income.

We also simulate larger shocks, which are more relevant for policymakers, and we find that a 33% reduction in trade costs results in a real income gain of 26%, with only 4% attributed to the mechanical effect. These results confirm that trade liberalization, by moving resources away from less distorted informal firms toward the initially more distorted formal sector, is able to substantially improve allocative efficiency. Although the impacts on the total number of informal workers are economically important, they are not enough to significantly curb informality. Specifically, as trade costs decline by 33%, the imports-to-GDP ratio increases from 7% to 30%, a substantial fourfold

---

<sup>3</sup>This decomposition relies on a first-order approximation of the changes in real income, following [Baqae and Farhi \(2020\)](#) and [Atkin and Donaldson \(2022\)](#).

<sup>4</sup>We define “measured aggregate total factor productivity” as the average of firm-level TFP weighted by employment, similarly to [Olley and Pakes \(1996\)](#) and [Pavcnik \(2002\)](#).

increase.<sup>5</sup> However, the share of informal workers declines much more modestly, from 51% to 46%.<sup>6</sup> This result is consistent with the casual observation that the informal sector has not substantially shrunk in middle-income economies despite the large-scale liberalization episodes these experienced in the 1980s and 1990s (see, for example, [World Bank, 2019](#)).

We investigate the extent to which the aggregate consequences of trade shocks we document can be attributed to the prevalence of informality within the economy. To this end, we compare the impact of a trade shock in the benchmark economy, representing the status quo in the data, to the outcomes in various scenarios in which we gradually increase government enforcement so that the size of the informal sector slowly declines, including the extreme case of perfect enforcement with zero informality. In all these counterfactual simulations, we hold the structural parameters fixed at the values estimated for the benchmark model and simply gradually increase the costs faced by informal firms. We find that the gains from trade, as measured by the increase in real income, are larger when enforcement is less strict (i.e., when there is more informality). This result is driven by changes in the magnitude of the reallocation effect, as the mechanical effect is approximately constant across scenarios. In the event of a small shock to trade costs, when we compare our extreme scenarios (benchmark and zero informality), the ratio between the magnitudes of the reallocation and mechanical effects declines from 4.7 to 2.0.<sup>7</sup> With a 33% reduction in trade costs, the total real income gains with no informality are less than half as large as those under our benchmark scenario (12% vs. 26%).

Our main conclusion is that an increase in globalization leads to a reallocation of labor and output from less distorted, initially overproducing informal firms to more distorted, initially underproducing formal firms and that this boosts the reallocation effect. In an economy with stricter enforcement, trade openness has less room to accomplish this reallocation as the informal sector is initially smaller, resulting in a relatively smaller magnitude of the reallocation effect.

Our paper contributes to a nascent literature on the interactions between trade and informality, e.g., [Goldberg and Pavcnik \(2003\)](#), [McCaig and Pavcnik \(2018\)](#), [Dix-Carneiro and Kovak \(2019\)](#), [Ponczek and Ulyssea \(2022\)](#). This literature empirically shows that the informal sector constitutes an important margin of adjustment to trade shocks. To date, this literature has relied on reduced-form empirical methods, resorting to difference-in-differences identification strategies. In particular, the last two papers focus on the unilateral Brazilian trade liberalization of the early 1990s and adopt a local labor markets approach. Their specifications compare labor market outcomes of a region facing a smaller (trade-induced) negative labor demand shock to those of a region facing a larger negative shock. Given that common impacts of trade are absorbed into the intercept, this type

---

<sup>5</sup>For comparison, the current imports-to-GDP ratios in the United Kingdom, South Africa, and Colombia are all approximately 30%.

<sup>6</sup>While the overall share of informality decreases by 5 percentage points, the fraction of informal employment within the manufacturing sector experiences a more substantial decline of 25%. This observation, coupled with our roundabout production structure and intersectoral linkages, helps explain the significant reallocation effect that we uncover.

<sup>7</sup>The reallocation effect is still important when informality is eradicated. As we elaborate in section 3.9, our model features other distortions, so that the no-informality equilibrium still exhibits inefficiencies.



of specification only isolates the impacts of negative labor demand shocks. From our model’s perspective, this involves comparing the benchmark economy with an economy experiencing a negative labor demand shock.

We examine how the predictions of our model compare to the findings of [Dix-Carneiro and Kovak \(2019\)](#) and [Ponczek and Ulyssea \(2022\)](#). Specifically, we simulate the impacts of negative labor demand shocks on unemployment, informality, and real income. The outcomes of our counterfactual experiments align with the findings that the informal sector acts as an “unemployment buffer” during periods of adverse labor market conditions. Notably, while the impact on unemployment remains relatively small, if not slightly negative, there is a significant expansion in the informal sector, consistent with the findings of [Dix-Carneiro and Kovak \(2019\)](#). Furthermore, [Ponczek and Ulyssea \(2022\)](#) show that the impact of negative shocks on unemployment (informality) is larger (smaller) in locations where the government more strictly monitors compliance with labor market regulations. Once again, our model’s predictions align with these findings, further validating the idea from both papers that the informal sector helps mitigate increases in unemployment during economic downturns.

However, we show that the aggregate real income losses from these negative labor demand shocks are larger in the benchmark economy with a larger informal sector than in the counterfactual economies with stricter enforcement and lower informality. Hence, the informal sector seems to serve as an unemployment buffer, but not as an aggregate real income buffer. The reason behind this result is that negative shocks push resources away from the more distorted formal sector toward the less productive and less distorted informal sector. This leads to a reallocation effect that is considerably more negative in the economy with a larger informal sector than in an economy in which the government more strictly monitors informal firms.

Given the scarcity of data on the informal sector, most empirical and quantitative studies on the impacts of trade have to focus on formal sector firms. In our final analysis, we investigate how incorporating the informal sector into these studies affects key insights and findings. We show that the standard measured aggregate productivity gains estimated from data on only the formal sector can substantially understate total aggregate productivity gains, which include the reallocation from low-productivity firms in the informal sector to high-productivity firms in the formal sector. Next, we revisit the effects of trade on cross-firm wage inequality. Prominent studies on this topic include [Coşar et al. \(2016\)](#) on Colombia and [Helpman et al. \(2017\)](#) on Brazil. Both papers center their analyses on data on formal manufacturing firms and find, from the lens of their quantitative structural models, that increases in openness were associated with increases in wage inequality across firms, driven by the behavior of the exporter wage premium and reallocation of labor from nonexporters to exporters. We replicate their finding that trade liberalization increases wage inequality in the formal manufacturing sector. However, we show that this result is reversed once one accounts for inequality in the whole economy, including the informal sector. Even though we account for only for between-firm wage inequality, our finding is particularly interesting in light of the recent literature that emphasizes the prominence of between-firm inequality in explaining

overall trends in wage inequality in different countries (e.g., [Card et al., 2013, 2016](#); [Song et al., 2019](#)), including Brazil ([Alvarez et al., 2018](#)).

The remainder of the paper is organized as follows. Section 2 describes the data we use in the analysis and presents five stylized facts on formal and informal workers and firms that we use to motivate our theoretical framework. Section 3 outlines our model. Section 4 details the estimation procedure, discusses identification and shows how the model fits key aspects of the data. Section 5 shows our counterfactual experiments, and section 6 presents our main takeaways.

## 2 Five Facts on Formal and Informal Firms and Workers in Brazil

We begin by highlighting five important facts about formal and informal workers and firms in Brazil that inform the selection of our model ingredients.<sup>8</sup> In anticipation of the notation introduced in the next section, we represent the manufacturing sector as  $C$  and services as  $S$ .

We make use of seven datasets containing information on formal and informal firms and their workers. An overview of these datasets and their main features is provided in Table 1. A key source of information on formal sector firms and workers is the *Relação Anual de Informações Sociais* (RAIS), which is a matched employer-employee dataset assembled by the Brazilian Ministry of Labor every year since 1976. RAIS is a high-quality panel that contains the universe of formal firms and workers.<sup>9</sup> With these data, we can provide a detailed cross-sectional picture of the formal labor market in the  $C$  and  $S$  sectors, and can generate important longitudinal statistics such as firm-level turnover and exit rates.

We also make use of three firm-level surveys conducted by the Brazilian National Statistics Agency (*Instituto Brasileiro de Geografia e Estatística*, IBGE) which cover the *formal* manufacturing, retail and service sectors: *Pesquisa Industrial Anual* (PIA), *Pesquisa Anual de Comércio* (PAC), and *Pesquisa Anual de Serviços* (PAS), respectively. These surveys collect detailed information on firms' revenues and inputs and combine a census of firms above a certain size threshold with a representative sample of smaller firms. Hence, longitudinal statistics can be computed for firms surveyed in the census. We identify exporters in RAIS and PIA by merging these datasets with administrative customs records from the *Secretaria de Comércio Exterior* (SECEX).

These five datasets provide a comprehensive view of the formal sector but omit information on the informal sector. Therefore, we use two additional sources providing data on informal firms and workers. The first is the *Pesquisa de Economia Informal Urbana* (ECINF), a cross-sectional survey collected by IBGE in 2003 and designed to be representative of the universe of *urban* firms with up to five employees (both formal and informal).<sup>10</sup> It is a matched employer-employee dataset that contains information on entrepreneurs, their businesses and their individual employees. Firms

<sup>8</sup>See [Ulyssea \(2020\)](#) for a more extensive discussion of stylized facts related to informality.

<sup>9</sup>The RAIS dataset has been increasingly used in different applications. For recent examples see [Dix-Carneiro \(2014\)](#), [Dix-Carneiro and Kovak \(2017\)](#), [Helpman et al. \(2017\)](#), and [Ulyssea \(2018\)](#), among others.

<sup>10</sup>Although a few firms in the dataset have more than five employees, we restrict our attention to those with five employees or fewer so that our sample is consistent with the population that the survey targets.

Table 1: Summary of Datasets

Dataset	Source	Description
Relação Anual de Informações Sociais <i>RAIS</i> Years: 2003–2005	Ministry of Labor	Administrative matched employer-employee dataset. Covers all formal firms and workers. Detailed information on firms and workers, but no information on firm-level revenues, capital and expenditures with intermediate inputs.
Pesquisa Industrial Anual <i>PIA</i> Years: 2003, 2004	IBGE	Survey data on Manufacturing firms. Firm-level information such as revenues, capital, investment, expenditures with intermediate inputs, employment. Covers all firms with 30 employees or more; random sample of smaller firms.
Pesquisa Anual dos Serviços <i>PAS</i> Years: 2003, 2004	IBGE	Survey data on Service-sector firms. Firm-level information such as revenues, capital, investment, expenditures with intermediate inputs, employment. Covers all firms with 20 employees or more; random sample of smaller firms.
Pesquisa Anual do Comércio <i>PAC</i> Years: 2003, 2004	IBGE	Survey data on Retail and Commerce firms. Firm-level information such as revenues, capital, investment, expenditures with intermediate inputs, employment. Covers all firms with 20 employees or more; random sample of smaller firms.
Secretaria de Comércio Exterior <i>SECEX</i> Year: 2003	Ministry of Industry, Foreign Trade and Services	Administrative customs data. Export and import values at the firm level.
Economia Informal Urbana <i>ECINF</i> Year: 2003	IBGE	Survey data. Matched employer-employee dataset with detailed information on formal and informal firms and their workers. Representative sample of small businesses (firms with 5 employees or less). Information on formal status of the firm and its workers.
Pesquisa Mensal de Emprego <i>PME</i> Year: 2003	IBGE	Survey data. Rotating panel of households that covers the 6 main metropolitan areas in Brazil.

are directly asked whether they are registered with the tax authorities and whether each of their workers has a formal labor contract. It is therefore possible to directly observe both firms’ and workers’ formal statuses. Given that the formality/informality statuses are self-reported, one might be concerned about measurement error and underreporting. However, IBGE has a long tradition of accurately measuring labor informality, and has very strict confidentiality clauses. The information that they collect cannot be used for auditing purposes by other government branches, in particular those responsible for enforcing the relevant laws and regulations. These characteristics, together with the high levels of informality observed in the data, make us confident that respondents are not systematically underreporting their informality status.<sup>11</sup>

Finally, we draw from the Monthly Employment Survey (*Pesquisa Mensal de Emprego*, PME) to obtain information on worker allocations and labor market flows. This is a rotating panel with a design similar to that of the Current Population Survey in the United States. The survey covers the six main metropolitan areas in Brazil and contains detailed information on individuals’ sociodemographic characteristics and labor market outcomes, including informal employment status. We define a worker to be informal if she does not hold a formal labor contract—this variable is

<sup>11</sup> Additionally, [Ulyssea \(2018\)](#) shows that the ECINF survey reproduces very well statistics from RAIS on all the dimensions that are common to both datasets (e.g., size and sectoral distributions), which is reassuring with respect to ECINF’s quality.

explicitly recorded in PME. If a worker is self-employed, she is also treated as an informal worker. Similarly, we treat self-employed workers in ECINF as informal firms employing one worker. We exploit the panel structure of PME to estimate one-year labor market transitions between formal employment, informal employment (in both  $C$  and  $S$  sectors) and unemployment status.<sup>12</sup>

We need to impose a few restrictions to make these seven datasets consistent with each other. First, because PME covers only six metropolitan regions, we restrict our samples to these regions whenever possible. Given the focus on metropolitan regions in PME and urban firms in ECINF, we remove firms and workers in the agriculture, mining, coal, and oil and gas industries.<sup>13</sup> Second, we restrict our attention to data from 2003, as ECINF is available only for that year. Whenever we need to compute dynamic statistics, we also employ data from 2004. Finally, we exclude firms in the public sector. Our data appendix provides additional details regarding data treatment—see Online Appendix L.

Having presented the various sources of data used in this paper, we are now ready to introduce the aforementioned five facts on formal and informal firms and workers in Brazil. These facts guide the construction of our model and are targeted in our estimation procedure.

***Fact 1: Approximately 50% of the Brazilian labor force is informally employed. Transitions from unemployment to an informal job are at least twice as likely as transitions from unemployment to a formal job.***

Table 2 uses data from PME to establish that 48.2% of all workers are informally employed. If we break down informality rates by sector, we find that 35.6% of  $C$  sector workers are informal, compared to 51.2% in the  $S$  sector. Moreover, transition rates from unemployment to informal jobs are twice as likely as those to formal jobs (45.3% compared to 21.1%).

Table 2: Employment Shares and Transition Rates

	Share of Workers	Transition Rates From Unemp.
Informal Manufacturing ( $Ci$ )	0.059	0.064
Formal Manufacturing ( $Cf$ )	0.106	0.050
Informal Services ( $Si$ )	0.351	0.389
Formal Services ( $Sf$ )	0.334	0.161
Unemployment	0.150	0.336
Share of Informal Employment	0.482	
Transition Rate from Unemp. to Informal Employment	0.453	
to Formal Employment	0.211	

Data source: 2003 PME.

***Fact 2: The probability that a firm is informal declines sharply with its employment size.***

<sup>12</sup>To minimize the effects of attrition in PME, we measure one-year transitions by annualizing four-month transitions.

<sup>13</sup>Incorporating the agriculture and mining sectors into the model would be desirable. These sectors play an important role in Brazil's international trade and they feature high prevalence of informality. Not being able to include them is an important limitation of our analysis.

We estimate, for each sector, regressions relating a firm-level informal status indicator to the firm’s number of employees. Table 3 uses data from ECINF to show that the fraction of informal firms rapidly declines with employment size.

Table 3: Firm-Level Informality Status vs. Firm-Level Employment

	Dep. Variable: <i>Informal Status Indicator<sub>i</sub></i>	
	<i>C</i> sector	<i>S</i> sector
Intercept	1.135 (0.028)	1.130 (0.012)
$\ell_i$	-0.179 (0.025)	-0.204 (0.009)
Observations	1,194	7,273

Data source: 2003 ECINF. Standard errors in parentheses.

***Fact 3: Informal firms have, on average, lower revenue per worker than formal firms.***

It has been widely documented that average worker productivity across firms is substantially lower in the informal sector than in the formal sector.<sup>14</sup> Our datasets confirm this insight. Note that the sample of informal firms from ECINF cannot be directly compared with the samples from PIA, PAS, and PAC as ECINF is designed to cover firms with at most five employees. However, we can compare firm-level labor productivity (measured as revenue per worker) across formal and informal firms by estimating linear regressions relating firm-level revenues to their employment size. The idea is that once we condition on employment size, the revenues of informal firms in ECINF can be compared to those of equal-sized formal firms in PIA, PAS, and PAC. The linear regression results reported in Panel A of Table 4 imply that, among *C* sector firms of size one, formal firms are on average more productive than informal ones by 1.73 log-points. For *C* sector firms of size five, this difference shrinks to 1.18 log-points. Similarly, for *S* sector firms of size one, formal firms are, on average, 1.18 log-points more productive than informal ones. Finally, for *S* sector firms of size five, this difference declines to 0.46 points.

***Fact 4: The average informal worker is paid lower wages than the average formal worker.***

It is a well-documented fact that average wages in the informal sector are lower than those in the formal sector. Although RAIS provides information on firm-level wages for the population of formal firms, we do not have firm-level wages for the population of informal firms. ECINF contains wages for informal workers but covers only firms with up to five employees. We compare wages across datasets by regressing log wages on log size for each type of firm in the two sectors (*C* and

<sup>14</sup>For example, refer to La Porta and Shleifer (2008), Meghir et al. (2015), and Ulyssea (2018). In this literature, worker productivity differences across firms are commonly measured by value added or revenue per worker. Whether this measure accurately reflects or is highly correlated with TFP at the firm level depends on the framework considered by the researcher. For example, in a standard Melitz model with heterogeneous firm-level TFP but no fixed costs of production, revenue per worker is constant across firms in equilibrium. However, the model that we discuss later in this paper, which incorporates labor market frictions and convex hiring costs, implies that firm-level productivity is predicted with high accuracy by revenue per worker. See Online Appendix P.

Table 4: Firm-Level log Revenue per Worker and log Wages *vs.* log Employment

Sector / Firm Type	A. Dep. Variable: $\log(\text{Revenue}_i/\ell_i)$				B. Dep. Variable: $\log(\text{wage}_i)$			
	<i>Cf</i>	<i>Sf</i>	<i>Ci</i>	<i>Si</i>	<i>Cf</i>	<i>Sf</i>	<i>Ci</i>	<i>Si</i>
Intercept	10.118 (0.013)	10.004 (0.005)	8.391 (0.037)	8.825 (0.015)	8.509 (0.006)	8.436 (0.002)	8.013 (0.033)	8.417 (0.014)
$\log(\ell_i)$	0.000 (0.005)	-0.128 (0.003)	0.342 (0.114)	0.321 (0.050)	0.117 (0.003)	0.105 (0.001)	0.292 (0.103)	0.231 (0.048)
<i>Exporter<sub>i</sub></i> (Dummy)	1.462 (0.021)				0.462 (0.014)			
Observations	16,986	43,861	1,070	6,202	20,075	145,981	1,071	6,205
Dataset	PIA + SECEX	PAS + PAC	ECINF	ECINF	RAIS + SECEX	RAIS	ECINF	ECINF

Standard errors in parentheses.

*S*). Panel B of Table 4 shows that once we control for size, we see an important informal wage penalty in the *C* sector but not in the *S* sector. Formal *C* sector firms of size one tend to pay wages that are 0.5 log points larger than those paid by their informal counterparts. For firms of size five this difference declines to 0.21 log points. On the other hand, *S* sector firms of size one pay approximately the same wages as their informal counterparts. Interestingly, if we condition on firms of size five, the second and fourth columns of Panel B of Table 4 imply that informal firms pay, on average, wages that are larger than their formal counterparts by 0.18 log points. Nonetheless, because informal firms tend to be considerably smaller than formal ones, an *average* informal worker receives a lower wage than an average formal worker.

***Fact 5: Firm-level labor turnover tends to decline with firm-level employment size. However, conditional on size, exporters tend to have higher turnover.***

In the context of Colombia, Coşar et al. (2016) show that labor turnover tends to decline with firm size but that, conditional on employment size, turnover tends to be larger for exporters. These relationships, replicated in Table 5 for Brazil, are potentially important for our quantitative analyses, as they establish a connection between trade openness and aggregate labor turnover. This link entails the reallocation of resources toward larger firms (which tend to exhibit lower turnover) and exporters (who, in contrast, tend to exhibit larger employment volatility). In a model incorporating search frictions and unemployment, turnover significantly influences the unemployment rate. Given Fact 1, these dynamics can, in turn, have substantial impacts on informality.

### 3 Model

#### 3.1 Motivating the Model's Ingredients

Before we start the exposition of our quantitative framework, we motivate our selection of its main ingredients, based on the five facts documented in the previous section. We begin with the premise that the existence of the informal sector is attributed to burdensome taxes and labor market regulations, which are imperfectly enforced by the government. While the government, in

Table 5: Turnover, Firm Size and Export Status

	Dep. Variable: $Turnover_i$	
	$C$ sector	$S$ sector
Intercept	0.741 (0.008)	0.645 (0.003)
$\log(\ell_i)$	-0.126 (0.003)	-0.096 (0.002)
$Exporter_i$ (Dummy)	0.071 (0.019)	
Observations	20,342	147,936

Data Sources: 2003 and 2004 RAIS and 2003 SECEX. Turnover of firm  $i$  between 2003 and 2004 measured as  $Turnover_i = \frac{|\ell_{i,2004} - \ell_{i,2003}|}{0.5 \times (\ell_{i,2004} + \ell_{i,2003})}$ .  
Standard errors in parentheses.

practice, imposes costs on informal firms through fines or the threat of closure, these measures are not applied uniformly across all firms. This lack of strict enforcement creates incentives for certain firms to operate under the radar, evading taxes and avoiding compliance with labor market regulations. In this context, Fact 2 in section 2 suggests that firms perceive a significant increase in the costs of informality as they grow. This increase is presumably linked to the heightened visibility of larger firms, leading to a larger probability of their receiving fines from the government.

Turning to other crucial components of the model, we note that Facts 1 and 5 require a model that includes unemployment and accommodates firm-level dynamics, resulting in worker turnover. Modeling unemployment is important in light of the empirical findings of [Dix-Carneiro and Kovak \(2019\)](#) and [Ponczek and Ulyssea \(2022\)](#), who document how trade-induced local labor demand shocks impact informality and unemployment.<sup>15</sup> It is our objective to make the model’s predictions align with these documented facts. In addition, firm-level dynamics and endogenous worker turnover are crucial for modeling the consequences of firing costs—a significant labor market regulation present in many Latin American countries, often blamed for the high levels of informality in the region (e.g., [Heckman and Pagés, 2000](#)).

To generate Fact 3, we need a model featuring firms that are heterogeneous in productivity. Furthermore, we allow firms to select their status (formal or informal), allowing for an endogenous response of the size of the informal sector to changes in the economic environment. Given Fact 2, smaller and less productive firms tend to choose to operate in the informal sector. This observation implies that, in equilibrium, less productive firms are subject to fewer distortions in their decisions from taxes and regulations than are more productive firms. This unequal exposure to domestic distortions can lead to considerable misallocation of resources. More precisely, higher-productivity formal firms face larger distortions, and, as a result, tend to underproduce relative to the social optimum. Meanwhile, lower-productivity informal firms face fewer distortions and tend

<sup>15</sup>In the context of the Brazilian trade liberalization of the early 1990s, [Dix-Carneiro and Kovak \(2019\)](#) show that (permanent) trade-induced negative labor demand shocks lead to long-run increases in informal employment but have no impact on unemployment. [Ponczek and Ulyssea \(2022\)](#) revisit these results and investigate the impact of the same trade-induced negative labor demand shocks on informality and unemployment across regions that enforce labor market regulations in varying degrees. They confirm that the informal sector helps mitigate the impact of adverse labor market conditions on unemployment.



to overproduce.

We next move on to Fact 4, to account for which the model must incorporate labor market frictions and equilibrium wage dispersion. The latter is particularly important for modeling the consequences of a minimum wage, another potentially significant labor market regulation influencing the incentives for firms to operate in the informal sector.

Finally, it is a priori important to model both the manufacturing and services sectors for two reasons. The first is the significant prevalence of informality in services (Table 2). The second is that service sector firms are also affected by trade through intersectoral linkages with manufacturing. This creates a potentially important channel linking trade to the aggregate level of informality.

Given the motivation for the model’s ingredients, we start by considering a closed economy setup in sections 3.2 through 3.5 and then extend the model to an open economy in section 3.6. The equilibrium conditions are outlined in section 3.7. In section 3.8, we delve into the mechanisms through which trade can impact informality, and section 3.9 wraps up with a final discussion of the model.

### 3.2 Consumers

The economy is populated by homogeneous, infinitely lived workers-consumers. Individuals derive utility from a final good whose production will be described shortly. Preferences are given by:

$$U = \sum_{t=1}^{\infty} \frac{Y_t}{(1+r)^t}, \quad (1)$$

where  $Y_t$  is the total amount of the final good consumed at time  $t$  and  $\frac{1}{1+r}$  is the discount factor. We denote the price of the final good at time  $t$  by  $P_{Y_t}$ .

### 3.3 Technology and Firms

The economy is comprised of two broad sectors: manufacturing ( $C$ ) and services ( $S$ ). Four types of goods are available: (a) the final consumption good, (b) sector-specific intermediate differentiated varieties, (c) sector-specific composite goods, and (d) sector-specific intermediate inputs. Sector-specific intermediate differentiated varieties (b) are aggregated into the sector-specific composite goods (c). In turn, these sector-specific composite goods are used in two ways: to produce the final good (a) and to produce the sector-specific intermediate inputs (d).<sup>16</sup> The sector-specific intermediate inputs (d) are finally bundled with labor to produce the intermediate differentiated varieties (b).<sup>17</sup>

We now describe the technologies available to produce these different types of goods. The final good, which is used solely for final consumption, is assembled by a perfectly competitive producer.

<sup>16</sup>Sector  $S$  composite goods are also used in the form of fixed costs of production, fixed costs of exporting and hiring costs.

<sup>17</sup>When we discuss the open economy model in section 3.6, we maintain that the final good (a), the sector-specific composite goods (c), and the sector-specific intermediate goods (d) cannot be traded across borders. However, the intermediate differentiated varieties (b) are allowed to be traded subject to trade costs.



Its output at time  $t$ , denoted by  $Y_t$ , is given by a Cobb-Douglas aggregate of the two sector-specific composite goods:

$$Y_t = \widehat{C}_t^\zeta \widehat{S}_t^{1-\zeta}, \quad (2)$$

where  $\zeta$  denotes the share of expenditures on the  $C$  sector composite good,  $\widehat{C}_t$  is the total usage of the sector  $C$  composite good, and  $\widehat{S}_t$  is the total usage of the sector  $S$  composite good.

In turn, sector-specific composite goods are produced by perfectly competitive producers operating in each sector  $k \in \{C, S\}$ . Total output of sector  $k$  is given by a constant elasticity of substitution (CES) aggregator over the output of a sector  $k$  specific continuum of differentiated varieties indexed by  $n \in [0, N_{kt}]$ , and where  $N_{kt}$  indicates the mass of available varieties in sector  $k$  at time  $t$ :

$$\mathcal{C}_t = \left( \int_0^{N_{Ct}} q_{Ct}(n)^{\frac{\sigma_C-1}{\sigma_C}} dn \right)^{\frac{\sigma_C}{\sigma_C-1}} \quad \mathcal{S}_t = \left( \int_0^{N_{St}} q_{St}(n)^{\frac{\sigma_S-1}{\sigma_S}} dn \right)^{\frac{\sigma_S}{\sigma_S-1}}. \quad (3)$$

In (3),  $\mathcal{C}_t$  denotes the total output of the sector  $C$  composite good,  $\mathcal{S}_t$  represents the total output of the sector  $S$  composite good,  $q_{kt}(n)$  is total usage of differentiated variety  $n$  from sector  $k$ , and  $\sigma_k > 1$  is the elasticity of substitution across varieties within sector  $k \in \{C, S\}$ .<sup>18</sup> These sector-specific composite goods are solely used as inputs for the production of the final good in (2) or as inputs for the production of sector-specific intermediate inputs, which we describe next. As we focus on steady-state equilibria, we henceforth drop the time subscript  $t$  for notational convenience.

The sector  $k$  specific intermediate input is also produced competitively and bundles the sector-specific composite goods determined by equation (3) as follows:

$$\iota_k \equiv \widetilde{C}^{\lambda_k} \widetilde{S}^{1-\lambda_k}, \quad (4)$$

where  $\lambda_k$  is the share of total payments to the sector  $C$  composite good,  $\widetilde{C}$  is the total usage of the sector  $C$  composite good, and  $\widetilde{S}$  is the total usage of the sector  $S$  composite good.

Units of differentiated variety  $n \in [0, N_k]$  for a particular sector  $k \in \{C, S\}$  are produced by firms that combine homogeneous labor with sector-specific intermediate inputs. Importantly, these are the firms in our model that map to the firms we observe in the data. From now on, whenever we refer to “firms,” we refer to these differentiated variety producers. Total output by a firm producing variety  $n$  in sector  $k$  with productivity  $z$  employing labor  $\ell$  and sector-specific intermediate inputs  $\iota_k$  is given by:

$$q_k(z, \ell, \iota_k) = z \ell^{\delta_k} \iota_k^{1-\delta_k}, \quad (5)$$

where  $\delta_k \in (0, 1)$ .

Guided by Facts 1 and 5, we generate firm-level dynamics and endogenous worker turnover by allowing firms’ idiosyncratic productivity  $z$  to evolve over time, following the AR(1) process below:

$$\ln z' = \rho_k \ln z + \sigma_k^z \varepsilon, \quad \rho_k \in (0, 1), \quad \varepsilon \sim N(0, 1), \quad (6)$$

where  $\sigma_k^z$  is the standard deviation of the stochastic shocks.<sup>19</sup>

<sup>18</sup>When we consider the open economy model, the set of available manufacturing varieties includes both domestically produced and foreign-produced varieties.

<sup>19</sup>This process is imposed to be the same across formal and informal firms within sectors  $C$  and  $S$ . Unfortunately,

Monopolistic competition among differentiated variety producers implies that gross revenues as a function of output  $q$  in sector  $k \in \{C, S\}$  are given by:

$$\tilde{R}_k(q) = \left( \frac{X_k}{P_k^{1-\sigma_k}} \right)^{\frac{1}{\sigma_k}} q^{\frac{\sigma_k-1}{\sigma_k}}, \quad (7)$$

where  $X_k$  is total expenditure on the sector  $k$  specific composite good,  $P_k = \left( \int_0^{N_k} p_k(n)^{1-\sigma_k} dn \right)^{\frac{1}{1-\sigma_k}}$  is the price of one unit of sector  $k$ 's composite good,  $k \in \{C, S\}$ , and  $p_k(n)$  is the price charged by firm  $n$  in sector  $k$ . To streamline the presentation, we henceforth refer to  $P_k$  as the price index of sector  $k$ .<sup>20</sup>

Aggregate expenditure on the sector  $C$  composite good is given by  $X_C = \zeta I + X_C^{int}$ , where  $I$  is aggregate income and  $X_C^{int}$  is the total expenditure from sector-specific intermediate input producers on the composite good from sector  $C$ . Similarly, aggregate expenditure on the sector  $S$  composite good is given by  $X_S = (1 - \zeta) I + X_S^{int} + E_S$ , where  $X_S^{int}$  is total expenditure from sector-specific intermediate input producers on the composite good from sector  $S$  and  $E_S$  represents expenditures on the  $S$  sector composite good made by firms in order to cover entry, hiring, fixed and export costs (which we discuss below). Aggregate income is determined by total wages, government transfers and aggregate firm profits.

Firms can freely adjust their intermediate input usage. Denote by  $\iota_k(z, \ell)$  the optimal intermediate input usage of a firm in sector  $k$  with productivity  $z$  and  $\ell$  workers. This firm's gross revenue can then be written as  $R_k(z, \ell) \equiv \tilde{R}_k \left( z \ell^{\delta_k} \iota_k(z, \ell)^{1-\delta_k} \right)$ . It is easy to show that expenditures on intermediates are proportional to gross revenues, resulting in the following expression for firm-level value added:

$$VA_k(z, \ell) = \frac{\sigma_k - (1 - \delta_k)(\sigma_k - 1)}{\sigma_k} R_k(z, \ell). \quad (8)$$

## Timing

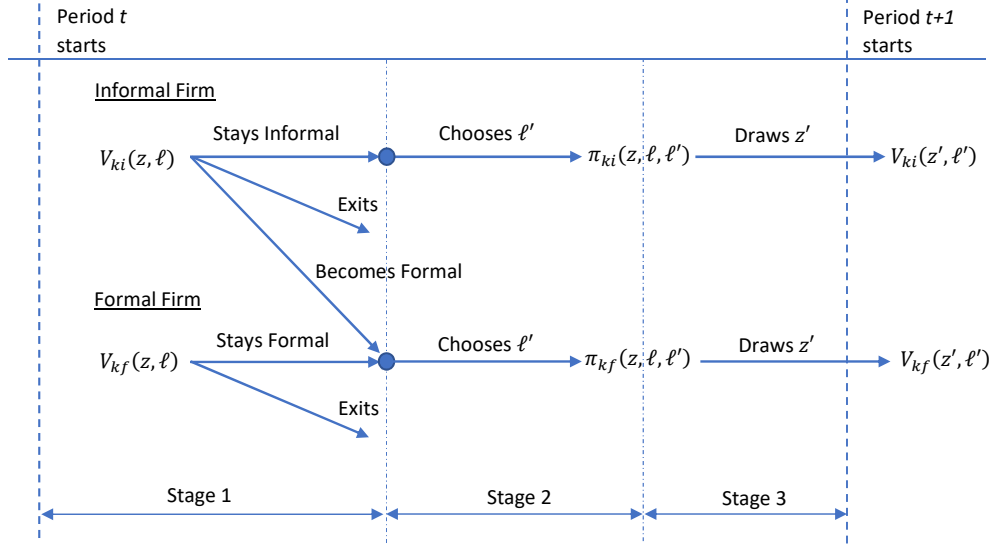
The timing of events is illustrated in Figure 1. Formal firms are indexed by  $f$  and informal firms by  $i$ . Consider an informal firm that starts period  $t$  with productivity  $z$  and employment level  $\ell$ . In the first stage, this firm may: (i) stay informal; (ii) exit, either as a result of an endogenous decision or because of an exogenous shock that occurs with probability  $\alpha_k$ ; (iii) become formal. In the second stage, the firm decides whether to adjust its workforce (up or down) to  $\ell'$  or not at all. Right after this decision the firm realizes profits and pays wages to its workers. In stage 3, the firm draws a new productivity value  $z'$  and starts period  $t + 1$  with state  $(z', \ell')$ . The timing and sequence of events for formal firms is the same as that for informal ones, except that we do not allow them to become informal.

---

we do not have longitudinal data on firms in the informal sector, so that this process cannot be separately identified for formal and informal firms.

<sup>20</sup>Given that  $P_k$  is the price of one unit of the sector  $k$  specific composite good,  $k \in \{C, S\}$ , the price of one unit of the final good is given by  $P_Y = \frac{P_C^\zeta P_S^{1-\zeta}}{\zeta^\zeta (1-\zeta)^{1-\zeta}}$ , and the price of one unit of the sector  $k$  specific intermediate input is given by  $P_k^m = \frac{P_C^{\lambda_k} P_S^{1-\lambda_k}}{\lambda_k^{\lambda_k} (1-\lambda_k)^{1-\lambda_k}}$ .

Figure 1: Diagram of Firms' Behavior



### Hiring and Firing Costs

Both formal and informal firms in the  $C$  and  $S$  sectors face hiring costs, which are incurred in units of the  $S$  sector composite good. We parameterize hiring costs as a function of employment levels  $\ell$  and number of posted vacancies  $v$ :

$$C_k^h(\ell, v) = \left( \frac{h_k}{\gamma_{k1}} \right) \left( \frac{v}{\ell^{\gamma_{k2}}} \right)^{\gamma_{k1}}, \quad (9)$$

where  $h_k, \gamma_{k1} > 1$  and  $\gamma_{k2} \in (0, 1)$  are parameters to be estimated. If  $\mu_{kj}^v$  is the probability of filling a vacancy faced by a firm of type  $j \in \{f, i\}$  in sector  $k \in \{C, S\}$ , then expanding from  $\ell$  to  $\ell'$  requires posting  $v = \frac{\ell' - \ell}{\mu_{kj}^v}$  vacancies.<sup>21</sup> The cost of expanding from  $\ell$  to  $\ell'$  workers for a firm of type  $j$  in sector  $k$  is therefore given by:

$$H_{kj}(\ell, \ell') = (\mu_{kj}^v)^{-\gamma_{k1}} \left( \frac{h_k}{\gamma_{k1}} \right) \left( \frac{\ell' - \ell}{\ell^{\gamma_{k2}}} \right)^{\gamma_{k1}}. \quad (10)$$

The value of  $\gamma_{k2}$  controls the extent to which firm-level growth rates in employment decline with size, a fact that we discussed in section 2 (Fact 5). The parameter  $\gamma_{k1}$  governs the convexity of the hiring function. Allowing for convexity is important for the model to be able to generate wage dispersion across firms, as emphasized in Fact 4. To build intuition for this fact, we momentarily abstract from dynamic considerations. In this case, the wage determination process that we discuss in section 3.5 implies that wages are proportional to value added per worker, which is—by virtue of our assumptions—proportional to marginal value added. Firms set marginal value added equal to the marginal cost of an additional worker. With linear hiring costs, the marginal cost is constant and equal across firms, so that wages will also be equalized across firms. In contrast, with convex hiring costs, the marginal cost of an additional worker is increasing in the growth of employment, so that expanding firms tend to pay higher wages.

<sup>21</sup>Note that the probability of filling a vacancy  $\mu_{kj}^v$  is an endogenous object that depends on the aggregate number of vacancies in each sector  $k \in \{C, S\}$  and firm type  $j \in \{f, i\}$  and on the mass of unemployed workers.

Firing costs are entirely driven by regulations and affect formal firms only. They take the linear form:

$$F(\ell, \ell') = \kappa(\ell - \ell'), \quad (11)$$

where  $\kappa > 0$  is the parameter governing the firing cost function. Consistent with the Brazilian labor market regulations, we assume that firing costs are equal across the  $C$  and  $S$  sectors. In our model, firing costs are collected by the government and are rebated back to consumers, while hiring costs are incurred in terms of the  $S$  sector composite good.

### Profit and Value Functions

Formal firms are subject to payroll and value-added taxes, firing costs and the minimum wage regulation. The profit function of a formal firm in sector  $k \in \{C, S\}$  is thus given by:

$$\pi_{kf}(z, \ell, \ell') = (1 - \tau_y) VA_k(z, \ell') - C_{kf}(z, \ell, \ell') - \bar{c}_k, \quad (12)$$

where  $\bar{c}_k$  denotes a per-period, fixed cost of operation, which we define in units of the  $S$  sector composite good;  $\tau_y$  is a value added tax, collected by the government and rebated lump-sum to consumers. Due to hiring and firing costs, the total cost function for a formal firm adjusting from  $\ell$  to  $\ell'$  workers is given by the following expression:

$$C_{kf}(z, \ell, \ell') = \begin{cases} (1 + \tau_w) \max\{w_{kf}(z, \ell'), \underline{w}\} \ell' + H_{kf}(\ell, \ell') & \text{if } \ell' > \ell \\ (1 + \tau_w) \max\{w_{kf}(z, \ell'), \underline{w}\} \ell' + \kappa(\ell - \ell') & \text{if } \ell' \leq \ell, \end{cases} \quad (13)$$

where the wage schedule  $w_{kf}(z, \ell')$  is the result of a bargaining problem between the firm and its workers that is detailed in section 3.5,  $\underline{w}$  denotes the minimum wage and  $\tau_w$  is the payroll tax, which is also assumed to be collected by the government and rebated lump-sum to consumers.

Since formal firms have to choose whether to stay or leave their industry, their value function is given by:

$$V_{kf}(z, \ell) = (1 - \alpha_k) \max \left\{ 0, \max_{\ell'} \left\{ \pi_{kf}(z, \ell, \ell') + \frac{1}{1+r} E_{z'|z} V_{kf}(z', \ell') \right\} \right\}, \quad (14)$$

where  $\alpha_k$  denotes the exogenous destruction probability that firms face every period for  $k = C, S$ . The solution of (14) leads to the employment policy function  $\ell' = L_{kf}(z, \ell)$  and to the vacancy posting policy function  $v_{kf}(z, \ell) = \frac{L_{kf}(z, \ell) - \ell}{\mu_{kf}^v} \times \mathcal{I}[L_{kf}(z, \ell) > \ell]$  (and to other policies such as the decisions to exit or stay active).

While informal firms do not incur any of the regulatory costs (taxes, minimum wages, firing costs), they do face a reduced-form expected cost of informality, which includes the probability of detection by the government and subsequent fines. It also includes a range of opportunity costs associated with informality such as scarce access to formal financial markets (e.g., credit lines), hampering the firms' ability to grow (Catão et al., 2009). As firms grow, they become more visible to the government and therefore face a higher probability of being inspected, which can entail costs in the form of fines or bribes or can lead to the firm's operations being shut down. Therefore, motivated by Fact 2, we allow the expected cost of informality as a fraction of revenues,  $p_{ki}$ , to

depend on the firm's size  $\ell'$ . Thus, the profit function of an informal firm is given by:

$$\pi_{ki}(z, \ell, \ell') = VA_k(z, \ell') - K^{inf}(z, \ell') - C_{ki}(z, \ell, \ell') - \bar{c}_k, \quad (15)$$

where  $K^{inf}(z, \ell') \equiv p_{ki}(\ell') R_k(z, \ell')$  are the expected costs associated with informality, which we assume proportional to gross revenues. We work with the following simple specification:

$$p_{ki}(\ell') = \tilde{a}_k \exp \left\{ \tilde{b}_k (\ell' - 1) \right\}. \quad (16)$$

The costs of informality incurred by firms are assumed to be rebated lump-sum to consumers.

Since informal firms are not subject to firing costs or other regulations, their cost function is given by:

$$C_{ki}(z, \ell, \ell') = \begin{cases} w_{ki}(z, \ell') \ell' + H_{ki}(\ell, \ell') & \text{if } \ell' > \ell \\ w_{ki}(z, \ell') \ell' & \text{if } \ell' \leq \ell, \end{cases} \quad (17)$$

where  $w_{ki}(z, \ell')$  denotes the wage paid by an informal firm with productivity  $z$  and size  $\ell'$ . The value functions of informal firms are similar to those of formal ones, except that they have the additional option of formalizing their businesses:

$$V_{ki}(z, \ell) = (1 - \alpha_k) \max \left\{ \begin{array}{l} 0, \max_{\ell'} \left\{ \pi_{ki}(z, \ell, \ell') + \frac{1}{1+r} E_{z'|z} V_{ki}(z', \ell') \right\}, \\ \max_{\ell'} \left\{ \pi_{kf}(z, \ell, \ell') + \frac{1}{1+r} E_{z'|z} V_{kf}(z', \ell') \right\} \end{array} \right\}. \quad (18)$$

The solution of (18) leads to the employment policy function  $\ell' = L_{ki}(z, \ell)$  and to the vacancy posting policy function  $v_{ki}(z, \ell) = \frac{L_{ki}(z, \ell) - \ell}{\mu_{ki}^v} \times \mathcal{I}[L_{ki}(z, \ell) > \ell]$  (and to other policies such as the decisions to exit, formalize their status or remain informal).<sup>22</sup>

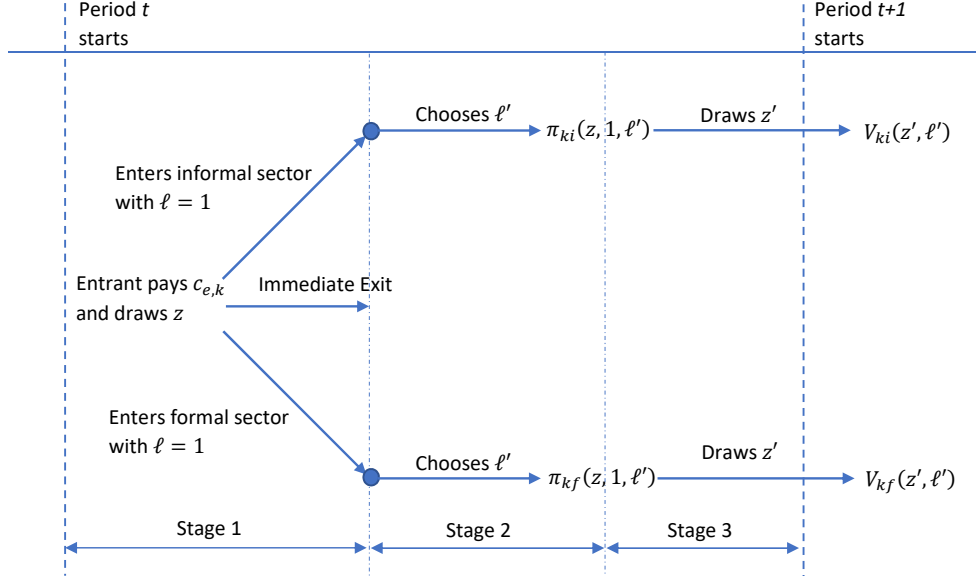
## Entry

Firm entry is illustrated in Figure 2. In each period there is a mass  $M_k$  of entrants into the  $C$  and  $S$  sectors. In the first stage within the period, entrants observe their productivity  $z$ —drawn from the ergodic distribution  $g_k^e$  implied by (6)—after incurring a sunk cost  $c_{e,k}$  of entry into sector  $k$ . Based on this productivity draw, the entering firm chooses to be formal or informal or to exit immediately. Formal and informal entrants start their first period with workforce  $\ell = 1$ , whose recruitment cost is subsumed in  $c_{e,k}$ . Following entry, in stage 2, the firm decides to adjust its labor force to  $\ell'$  just before the production stage. It then behaves as an incumbent, drawing productivity  $z'$  for the next period right after production (stage 3). The value functions for entrants in either sector are given by:

$$V_{kj}^e(z) = \max_{\ell'} \left\{ \pi_{kj}(z, 1, \ell') + \frac{1}{1+r} E_{z'|z} V_{kj}(z', \ell') \right\}, \quad (19)$$

<sup>22</sup>This setup, where informal firms can transition to formal status but formal firms are not allowed to shift to informality, leads to overlapping distributions of productivity  $z$  within each sector  $k \in \{C, S\}$  in equilibrium. This is an important feature of the data on formal and informal firms for many countries, as emphasized by Meghir et al. (2015). It is worth noting that once a firm becomes formal, it persists in operating formally as long as it remains profitable, even if its productivity drops below the threshold for entry into the formal sector.

Figure 2: Diagram of Entry Behavior



where  $j = i, f$ . The value at entry together with the entry conditions is defined by

$$V_k^e = E_z \max \{ V_{ki}^e(z), V_{kf}^e(z), 0 \}, \quad (20)$$

whose solution leads to the entry policy functions. Assuming there are positive masses of entrants in each sector, free entry dictates that:

$$V_k^e = c_{e,k}. \quad (21)$$

### 3.4 Labor Market Frictions

Motivated by Facts 1 and 4 documented in section 2, both formal and informal labor markets are characterized by search and matching frictions. These frictions prevent unemployed workers from immediately finding open vacancies and contribute to hiring costs, as highlighted in equation (10). We assume random search, and therefore all unemployed workers form a pool of individuals who randomly meet with formal or informal firms in one of the sectors  $k = C, S$ . Thus, formal and informal firms operating in the  $C$  and  $S$  sectors compete for workers in the labor market. Given the total number of vacancies posted in each sector and type of firm ( $v_{Cf}, v_{Ci}, v_{Sf}, v_{Si}$ ), and the mass of unemployed workers searching for jobs,  $L_u$ , the total number of matches formed is given by:

$$m(v_{Cf}, v_{Ci}, v_{Sf}, v_{Si}, L_u) = \phi \tilde{v}^\xi L_u^{1-\xi}, \quad (22)$$

where  $\tilde{v} = v_{Cf} + v_{Ci} + v_{Sf} + v_{Si}$  aggregates vacancies across sectors and types of firms,  $\phi > 0$ , and  $0 < \xi < 1$ . Matches are split across sectors and firm types in proportion to the number of vacancies posted, so that  $m_{kj} = \frac{v_{kj}}{\tilde{v}} \times m$  matches are formed with firms of type  $j$  in sector  $k$ . Thus, the probability of filling a vacancy ( $\mu_{kj}^v = \frac{m_{kj}}{v_{kj}}$ ) is independent of sector and firm type. We denote it

by  $\mu^v$ , and it is given by:

$$\mu^v = \phi \left( \frac{L_u}{\widetilde{v}} \right)^{1-\xi}. \quad (23)$$

This expression highlights that formal firms directly compete with informal ones in the labor market. Finally, unemployed workers face job-finding probabilities in each sector and firm type given by<sup>23</sup>:

$$\mu_{kj}^e \equiv \frac{m_{kj}}{L_u} = \frac{v_{kj}}{\widetilde{v}} \left( \frac{\phi}{(\mu^v)^\xi} \right)^{\frac{1}{1-\xi}}. \quad (24)$$

### 3.5 Wages

Wage setting takes place after hiring and fixed costs have been sunk and matching has taken place. We assume that a union engages in collective bargaining with the employer on behalf of the workers over the surplus of the match, determining a wage  $w_{kj}(z, \ell')$ . The latter depends on the firm's size and productivity.

The surpluses of a formal firm in sector  $k$  ( $S_{kf}^e$ ) and the union that it faces ( $S_{kf}^u$ ) are each defined as:

$$S_{kf}^e(z, \ell') = (1 - \tau_y) V A_k(z, \ell') - (1 + \tau_w) w_{kf}(z, \ell') \ell' + \frac{1}{1+r} E_{z'|z} V_{kf}(z', \ell'), \quad (25)$$

$$S_{kf}^u(z, \ell') = \left[ w_{kf}(z, \ell') + \frac{1}{1+r} J_{kf}^e(z, \ell') - \left( b_0 \times P_Y + b^u + \frac{1}{1+r} J^u \right) \right] \ell', \quad (26)$$

where  $b_0$  denotes the utility flow from being unemployed;  $P_Y$  is the price of one unit of the final good;  $b^u$  denotes the flow of unemployment insurance benefits, which are only received by formal workers;  $J^u$  is the expected present value of search; and  $J_{kf}^e(z, \ell')$  is the expected present value of a job in a formal firm in sector  $k$  with current productivity  $z$  and workforce  $\ell'$ —see Online Appendix I for its derivation.<sup>24</sup>

Let  $\beta$  be the parameter that drives workers' bargaining power. If the joint surplus of the firm and workers is positive, the outcome of bargaining is given by:

$$S_{kf}^u(z, \ell') = \beta (S_{kf}^e(z, \ell') + S_{kf}^u(z, \ell')). \quad (27)$$

Importantly, the overall surplus depends on the wage because of payroll taxes: in other words, the value of the surplus depends on how it is shared. This leads to a wage structure for formal workers defined by:

$$\begin{aligned} (1 + \beta \tau_w) w_{kf}^u(z, \ell') &= (1 - \beta) \left( b_0 \times P_Y + b^u + \frac{1}{1+r} (J^u - J_{kf}^e(z, \ell')) \right) \\ &+ \beta \left( (1 - \tau_y) \frac{V A_k(z, \ell')}{\ell'} + \frac{1}{1+r} E_{z'|z} \frac{V_{kf}(z', \ell')}{\ell'} \right). \end{aligned} \quad (28)$$

Thus, the bargained wage is proportional to a convex combination between the firm's value per worker and the worker's outside option net of the continuation value of the job to the worker,

<sup>23</sup>  $\mu_{kj}^e$  should be interpreted as the transition rate from unemployment to sector  $k$  and firm type  $j$ .

<sup>24</sup> Two observations about equations (25) and (26): (a) we assume that if all workers leave, the firm exits, and that hiring costs and fixed operating costs are already sunk at the bargaining process; and (b) note that  $b_0$  is multiplied by  $P_Y$  in (26), converting it into a nominal value.

$J_{kf}^e(z, \ell')$ .<sup>25</sup> If  $w_{kf}^u(z, \ell')$  leads to a negative union surplus then we set the wage equal to its reservation value, that is, the wage  $w_{kf}^{res}(z, \ell')$  that solves  $S_{kf}^u(z, \ell') = 0$ . Therefore the outcome of the Nash bargaining process leads to wage schedule:

$$w_{kf}(z, \ell') = \max \{w_{kf}^u(z, \ell'), w_{kf}^{res}(z, \ell')\}. \quad (29)$$

As highlighted in equation (13), formal firms cannot pay below the minimum wage, so that wages effectively paid are given by  $\max \{w_{kf}(z, \ell'), \underline{w}\}$ .

Wages in the informal sector are determined in a similar way. However, in the informal sector unemployment insurance benefits are not offered, taxes are not paid and firms face an expected cost of informality. Thus, the wage that informal firms pay their workers is given by:

$$w_{ki}(z, \ell') = \max \{w_{ki}^u(z, \ell'), w_{ki}^{res}(z, \ell')\}, \quad (30)$$

where

$$\begin{aligned} w_{ki}^u(z, \ell') = (1 - \beta) & \left( b_0 \times P_Y + \frac{1}{1 + r} (J^u - J_{ki}^e(z, \ell')) \right) \\ & + \beta \left( \left( 1 - \frac{\sigma_k}{\sigma_k - (1 - \delta_k)(\sigma_k - 1)} \right) p_{ki}(\ell') \frac{V A_k(z, \ell')}{\ell'} + \frac{1}{1 + r} E_{z'|z} \frac{V_{ki}(z', \ell')}{\ell'} \right), \end{aligned} \quad (31)$$

and  $J_{ki}^e(z, \ell')$  is the analogous to  $J_{kf}^e(z, \ell')$  in the informal sector, see Online Appendix I for details. Finally, the reservation wage,  $w_{ki}^{res}(z, \ell')$ , solves  $S_{ki}^u(z, \ell') = 0$ , where  $S_{ki}^u$  is the surplus of the union in the informal sector. This places a lower bound on wages in the informal sector.

### 3.6 Open Economy

We now extend the model to the open economy case. We assume that the home country is small relative to the rest of the world and therefore that foreign conditions do not react to the home country's policies. In the following analysis, we drop the formal/informal qualifier to simplify notation, as we assume throughout that informal firms cannot export.<sup>26</sup>

#### Price Indices and Aggregate Variables

Let  $N_{F,C}$  denote the measure of foreign varieties available to domestic consumers. Given the small open economy assumption, this variable is assumed to be fixed. Without loss of generality, we normalize the price index of free-on-board imports to be one, since foreign prices are exogenous to our model. Thus, the price index of imports in domestic currency becomes  $P_{F,C} = \epsilon \tau_a \tau_c$ , where  $\epsilon$  is the exchange rate,  $\tau_a - 1 > 0$  is the ad valorem tariff and  $\tau_c > 1$  the iceberg trade cost. The price index of domestically produced differentiated varieties  $n \in (N_{F,C}, N_C]$  is given by  $P_{H,C} = \left( \int_{N_{F,C}}^{N_C} p(n)^{1-\sigma_C} dn \right)^{\frac{1}{1-\sigma_C}}$ , and the price index for the composite  $C$  sector good is given by  $P_C = \left[ P_{H,C}^{1-\sigma_C} + P_{F,C}^{1-\sigma_C} \right]^{\frac{1}{1-\sigma_C}}$ .

<sup>25</sup>The factor of proportionality  $1 + \beta \tau_w$  highlights that workers also bear the cost of payroll taxes.

<sup>26</sup>This assumption comes from the fact that unregistered firms cannot undertake the necessary legal and bureaucratic procedures to export.



The domestic demand for varieties produced domestically is given by  $Q_{H,C}(n) = D_{H,C}p(n)^{-\sigma_C}$ , for  $n \in (N_{F,C}, N_C]$  with  $D_{H,C} \equiv \left(\frac{X_C}{P_C^{1-\sigma_C}}\right)$ ; and the domestic demand for foreign-produced varieties is given by  $Q_{H,C}(n) = D_{H,C}(\epsilon\tau_a\tau_c p^*(n))^{-\sigma_C}$ , for  $n \in [0, N_{F,C}]$ —where  $p^*(n)$  is the price of foreign variety  $n$  in foreign currency. Finally, foreign demand for domestically produced differentiated varieties is given by  $Q_{F,C}(n) = D_F^*(p_x^*(n))^{-\sigma_C}$ , for  $n \in (N_{F,C}, N_C]$ , where  $p_x^*(n)$  is the price of domestic variety  $n$  in the foreign country, denominated in foreign currency, and  $D_F^*$  is an exogenous foreign demand shifter. If  $\mathcal{I}_C^x(n)$  denotes an indicator function that equals one if variety  $n$  is exported, we have that the values of aggregate imports (before import tariffs) and exports are given by the following expressions:

$$Imports = \frac{D_{H,C}(\epsilon\tau_a\tau_c)^{1-\sigma_C}}{\tau_a} \quad \text{and} \quad Exports = D_F^* \epsilon \int_{N_{F,C}}^{N_C} \mathcal{I}_C^x(n) p_x^*(n)^{1-\sigma_C} dn. \quad (32)$$

## Exporters

Given the expression of foreign demand for home differentiated variety  $n$  just described,  $Q_{F,C}(n)$ , revenues from exports are given by  $\epsilon D_F^* \frac{1}{\sigma_C} (q_x/\tau_c)^{\frac{\sigma_C-1}{\sigma_C}}$ , where  $q_x$  is the total quantity exported. If a firm exports, it must decide which fraction  $\eta$  of its output to sell abroad. *Conditional on the firm's being an exporter*, total gross revenue for producing a total of  $q$  units and exporting a fraction  $\eta$  of this production is given by:  $\tilde{R}_C^x(q, \eta) = \exp(d_{H,C} + d_F(\eta)) q^{\frac{\sigma_C-1}{\sigma_C}}$ , where  $d_{H,C} = \ln\left(D_{H,C}^{\frac{1}{\sigma_C}}\right)$  and  $d_F(\eta) = \ln\left((1-\eta)^{\frac{\sigma_C-1}{\sigma_C}} + \epsilon\left(\frac{D_F^*}{D_{H,C}}\right)^{\frac{1}{\sigma_C}}\left(\frac{\eta}{\tau_c}\right)^{\frac{\sigma_C-1}{\sigma_C}}\right)$ . It is easy to verify that all exporters optimally decide to export the same fraction,  $\eta^o$ , of their production. In what follows, we will refer to  $d_F(\eta^o)$  as simply  $d_F$ .<sup>27</sup> Empirically,  $d_F$  is directly related to the fraction of gross revenues coming from exports among exporters, which is given by:

$$\frac{R_C^x(z, \ell') - R_C^{dom,x}(z, \ell')}{R_C^x(z, \ell')} = 1 - \exp(-\sigma_C \times d_F), \quad (33)$$

where  $R_C^x(z, \ell') \equiv \tilde{R}_C^x(q_C(z, \ell', \iota_C(z, \ell')), \eta^o)$  is the total gross revenue of an exporter with state  $(z, \ell')$  and  $R_C^{dom,x}(z, \ell')$  is the portion of an exporter's gross revenues coming from domestic sales.

The value added function for exporters takes the form:

$$VA_C^x(z, \ell') = \exp\left(\frac{\sigma_C}{\sigma_C - 1} \Lambda_C \times d_F\right) VA_C^d(z, \ell'), \quad (34)$$

where  $\Lambda_C \equiv \frac{\sigma_C-1}{\sigma_C-(1-\delta_C)(\sigma_C-1)}$  and  $VA_C^d$  is the value-added function for nonexporters. It follows that the export decision is given by:

$$\mathcal{I}_C^x(z, \ell') = \begin{cases} 1 & \text{if } VA_C^x(z, \ell') - f_x > VA_C^d(z, \ell') \\ 0 & \text{otherwise,} \end{cases} \quad (35)$$

where  $f_x > 0$  denotes the fixed cost of exporting, which is denominated in terms of the  $S$  sector composite good.

---

<sup>27</sup>When we substitute  $\eta^o$  into  $d_F(\eta)$ , we obtain  $d_F \equiv d_F(\eta^o) = \log\left(\left(1 + \frac{D_F^*}{D_{H,C}} \epsilon^{\sigma_C} \tau_c^{1-\sigma_C}\right)^{\frac{1}{\sigma_C}}\right)$ .

### 3.7 Equilibrium

We now summarize the equilibrium conditions below. Online Appendices A to I give further details.

1. Firms act optimally, make entry and exit decisions, and post vacancies according to equations (14), (18), (19) and (20). If entry is positive in sector  $k$ , the free-entry condition (21) holds with equality.
2. Wage schedules solve the bargaining problem between workers and the firm, as in equations (29) and (30).
3. Labor markets clear, that is, the sum of employment levels across sectors and firms types and the number of unemployed workers must be equal to the total labor force  $\bar{L}$ .
4. Product markets clear; the sum of expenditures on each sector-specific composite good, including consumption, intermediate goods, costs of entry, hiring, and export costs must add up to revenues in the sector (where relevant, this includes payment of tariffs).
5. Trade is balanced:  $Imports = Exports$ .
6. The government runs a balanced budget. All government revenues stemming from tax collection (including tariff revenues and fines on informal firms) and firing costs must exceed expenditures with unemployment benefits to all unemployed workers dismissed from formal employment. The budget surplus is directly rebated to consumers.
7. Aggregate income  $I$  is given by the sum of all wages and profits, plus the revenue from tariffs  $(\tau_a - 1) \times Imports$ , minus the total costs incurred by entering and hiring firms, fixed costs of operation and fixed costs of exporting.
8. We focus on steady-state equilibria, where the distributions of states  $(z, \ell)$ , by sector and firm type, and all aggregate variables remain constant. In particular, no sector can be expanding or contracting, which implies that: (i) the flow of workers out of unemployment and into the formal/informal and  $C/S$  sectors must be the same as the flow out of these sectors and into unemployment; (ii) the mass of firms entering the informal sector must be equal to the mass of informal firms that decide to exit or to formalize their businesses in either sector  $k \in \{C, S\}$ ; and (iii) the sum of the number of firms entering the formal sector and those formalizing their businesses must equal the mass of formal firms that decide to exit either sector  $k \in \{C, S\}$ .

### 3.8 Trade and Informality: Discussion of Mechanisms

Our model includes several channels through which trade can impact informality, pushing the response to changes in trade openness in different directions. The first set of mechanisms linking trade to informality is what we call “Melitz-type” mechanisms, which operate through various channels. In the subsequent discussion, we abuse language to some extent, as firms’ policy functions

depend on both their productivity  $z$  and workforce size  $\ell$ . However, for simplicity, we describe our mechanisms by focusing on selection based solely on productivity  $z$ .

Declines in import tariffs  $(\tau_a - 1)$  and iceberg trade costs  $(\tau_c)$  lead to an increase in the size of the foreign market perceived by  $C$  sector firms ( $d_F$ ). Consequently, initially higher-paying exporters experience greater profitability, prompting them to further raise wages and expand their workforce. Moreover, this shift encourages new firms to become high-paying exporters, further contributing to higher aggregate wages in sector  $C$ . These increased wages, in turn, elevate the value of search ( $J^u$ ), exerting upward pressure on wages across all firms and sectors in the economy, as in Melitz (2003). These higher wages tend to push unproductive informal firms out of the market, raising the productivity threshold for operation in the informal sector. All else equal, this movement tends to lead to a decline in informality. Similarly, low-productivity formal firms find incentives to exit the market as their labor costs increase, and the productivity threshold for new firm entry into the formal sector is raised. *Ceteris paribus*, this tends to lead to an increase in informality. Taken together, these impacts lead to ambiguous effects of trade openness on aggregate informality.

Following a decline in the iceberg trade cost or tariff, exporters not only experience increased profitability but also see a higher likelihood that any firm in sector  $C$  can eventually become an exporter. Firms, being forward looking, observe these changes, leading to an increased present value of formality compared to informality in sector  $C$ . This, in turn, encourages high-productivity informal firms to transition to the formal sector. As a result, the productivity threshold for operating in the formal sector tends to decrease. When all other factors held constant, this force tends to reduce informality.

In addition, reductions in tariffs and trade costs affect the price of the intermediate composites  $\iota_k$ , effectively altering optimal intermediate input usage by firms and, therefore, their workers' productivity. This effect on intermediate inputs amplifies the impact of tariff or trade cost reductions and is similar in spirit to the "magnification" effect highlighted in the work of Fieler et al. (2018) and Coşar et al. (2016) among others. A change in trade policy or other trade costs hence directly affect the decisions of firms, including their decisions to enter, exit, and produce formally or informally. Specifically, a reduction in trade barriers tends to make all firms more productive by improving access to cheaper intermediate goods. This encourages the most productive informal firms to grow and formalize. However, it can also lead to entry of lower-productivity informal firms, which have now become profitable. Furthermore, trade openness induces a reallocation of resources toward larger and more productive firms that export. This tends to further increase aggregate productivity and income, shifting aggregate demand up, generating incentives for productive informal firms to grow and formalize. The net effect of all these forces depends on the values of the parameters that we estimate.

The second mechanism linking openness to informality in our model is through its effect on unemployment. We show in section 2 that, in the Brazilian data, transitions from unemployment to informal employment are twice as large as transitions from unemployment to formal employment. Therefore, the channels in our model linking openness to unemployment have implications for the

relative importance of the informal sector.

Accordingly, we now turn to the mechanisms linking trade to unemployment in our setup. All else equal, equation (34) shows that exporters' value added is magnified relative to nonexporters', since  $\exp\left(\frac{\sigma_C}{\sigma_C-1}\Lambda_C \times d_F\right) > 1$ . This implies that exporters' value added, and therefore their hiring and firing decisions, are more sensitive to productivity shocks  $z$ . Therefore, as in Coşar et al. (2016), reducing trade costs produces two opposing forces: (i) there is reallocation of workers toward larger and higher-productivity firms, which tend to be more stable and have lower worker turnover (as they face larger costs of growing the workforce); and (ii) due to the term  $\exp\left(\frac{\sigma_C}{\sigma_C-1}\Lambda_C \times d_F\right)$ , both new and old exporters become more sensitive to idiosyncratic shocks, which tends to increase turnover. We follow Coşar et al. (2016) and refer to these two forces as the distribution effect and sensitivity effect, respectively. The bottom line is that increasing openness can increase or decrease labor turnover depending on which of the two effects dominates. Labor turnover is tightly linked to unemployment, as workers who are fired must spend at least one period in unemployment. As a result, increasing openness can lead to increases or decreases in unemployment. Moreover, as explained in the previous paragraph, these unemployment changes generate corresponding changes in the share of informality.

### 3.9 Discussion

The key premise of our model is that the Brazilian government imposes burdensome and distortive taxes and regulations on the country's firms. These taxes and regulations distort the decisions of compliant formal firms. However, because of imperfect enforcement, firms are not equally exposed to these distortions: informal firms bypass taxes and labor market regulations, but face size-dependent costs of informality.

This environment implies that low-productivity and small-scale informal firms tend to overproduce relative to what is socially optimal while high-productivity formal firms tend to underproduce. In addition, according to Fact 4, informal firms enjoy a cost advantage, sustaining lower wages, thereby further amplifying misallocation.

The model features other distortions such as (a) monopolistic competition; (b) markups that differ across the  $C$  and  $S$  sectors; (c) roundabout production amplifying the aforementioned distortions; (d) search frictions, unemployment and Nash bargaining over wages. These distortions are not at the center of the analysis as they are not directly related to the phenomenon of informality, but we nevertheless include them in the analysis to have a realistic framework for the quantitative exercises.

In particular, labor market frictions are important to generate the unemployment and informality impacts documented in previous work (e.g., Dix-Carneiro and Kovak, 2019; Ponczek and Ulyssea, 2022). Intersectoral linkages are important to the model as they allow us to link trade shocks to labor market conditions in services, a large sector with a high prevalence of informality. Finally, the monopolistic competition environment that we adopt is a standard framework for analyzing the impacts of trade liberalization with heterogeneous firms (e.g., Melitz, 2003).

## 4 Estimation

We now quantify the model outlined in section 3 in two steps. First, we fix a subset of parameters based on a combination of aggregate data, estimates from previous papers, and the statutory value of institutional parameters, such as value-added and payroll taxes. Next, we estimate the remaining model parameters using an Indirect Inference procedure, which allows us to combine information from the different data sources discussed in section 2. Section 4.1 describes how we determine the parameters fixed throughout the estimation procedure. Section 4.2 discusses the estimation procedure, and section 4.3 addresses identification. Finally, section 4.4 presents the estimation results and discusses model fit.

### 4.1 Fixed Parameters

We describe the relevant labor market regulations in Brazil in Online Appendix N. As has been extensively discussed in previous work (Heckman and Pagés, 2000; Gonzaga et al., 2003; Botero et al., 2004), not only these labor market regulations impose significant costs on employers, but they also involve a certain level of complexity. Therefore, we need to make simplifying assumptions to map them to our model’s payroll tax ( $\tau_w$ ), value-added tax ( $\tau_y$ ), firing cost ( $\kappa$ ), and unemployment insurance benefit ( $b_u$ ). We follow Ulyssea (2018) and set  $\tau_w$  so that it reflects the main taxes that are proportional to firms’ wage bill, namely, the employers’ social security contribution (20%), payroll tax (9%), and contributions to the severance indemnity fund, the *Fundo de Garantia do Tempo de Serviço* (FGTS; 8.5%). We combine two VAT-like taxes to calculate  $\tau_y$ : *Imposto sobre Produtos Industrializados* (IPI; 20%) and the Social Integration Program/Contribution for the Financing of Social Security PIS/COFINS (9.25%).<sup>28</sup>

Firing costs are set based on Heckman and Pagés (2000), who compute the cost of dismissing workers for several Latin American countries, including Brazil. Their calculation takes into account the specific features of dismissal costs that we review in Online Appendix N and shows that, on average, employers must pay approximately 1.9 months of wages to dismiss a worker. Considering that the *annual* average formal-sector wage in the 2003 RAIS data amounts to R\$10,565, we obtain a firing cost  $\kappa$  of R\$1,690 per worker. The minimum wage corresponds to the annualized value of the national monthly minimum wage in 2003:  $\underline{w} = \text{R}\$2,880$ . To compute unemployment insurance benefits, we assume that all workers receive the maximum duration of potential benefits (that is, 5 monthly payments). This figure is very close to both the mean and median of the duration of actually received benefits (Gerard and Gonzaga, 2021). Finally, we use the average monthly value of benefits paid in 2003, as reported by the Ministry of Labor: 1.37 times the minimum wage. This amounts to unemployment insurance benefits  $b_u = 1.37 \times 5$  times the *monthly* minimum wage = R\$1,644.

The share of final consumption expenditure on the sector  $C$  composite good,  $\zeta$ , and the share

---

<sup>28</sup>We exclude state-level value-added taxes because these vary greatly across states and there is a complex system of tax substitution across the production chain, which would be impossible to properly capture.

of sector  $k$ 's intermediate input payments to the sector  $C$  composite good,  $\lambda_k$ , are extracted from the 2000 and 2005 Brazilian National Accounts. We obtain  $\zeta = 0.283$ ,  $\lambda_C = 0.65$ , and  $\lambda_S = 0.29$ , suggesting that tariffs and iceberg trade costs can have a substantial effect on labor productivity in both sectors. The iceberg trade cost  $\tau_c = 2.4$  is obtained based on Head and Ries index (Head and Mayer, 2014) applied to 2003 trade flows between Brazil and the rest of the world, and the average import tariff  $\tau_a - 1 = 0.12$  comes from 2003 data in UNCTAD TRAINS.<sup>29</sup>

There are two sets of model parameters that are hard to identify given our data: the bargaining weight of workers  $\beta$  and the matching function parameters ( $\phi$  and  $\xi$ )—see Flinn (2006) for a discussion. We follow Mortensen and Pissarides (1999) and Ljungqvist and Sargent (2017) and impose symmetric bargaining, i.e.  $\beta = 0.5$ . We assume a matching function elasticity of  $\xi = 0.5$ , which is in the middle of the range surveyed by Petrongolo and Pissarides (2001), and  $\phi = 0.576$ , a choice we discuss in section 4.3. Table 6 summarizes the parameter values fixed throughout the estimation and their sources.

Table 6: Fixed Parameters

Parameter	Description	Source	Value
$\tau_c$	Iceberg Trade Cost	Head and Ries Index + 2003 WIOD	2.4
$\zeta$	Share of final expend. on $C$	IBGE National Accounts (2000/2005)	0.283
$\lambda_C$	Prod. Function (Equation (4))	IBGE National Accounts (2000/2005)	0.645
$\lambda_S$	Prod. Function (Equation (4))	IBGE National Accounts (2000/2005)	0.291
$r$	Interest rate	Ulyssea (2010)	0.08
$\tau_y$	Value Added Tax	Ulyssea (2018)	0.293
$\tau_w$	Payroll Tax	Ulyssea (2018)	0.375
$\tau_a - 1$	Import Tariff	UNCTAD TRAINS	0.12
$\kappa$	Firing Costs (in R\$)	Heckman and Pagés (2000)	1,956.7
$\underline{w}$	Min. Wage (in R\$)	Annualized 2003 value	2,880
$b_u$	Unemployment Benefit	$1.37 \times 5 = 6.85$ monthly Min. Wage	1,644
$\xi$	Matching Function	Petrongolo and Pissarides (2001)	0.5
$\phi$	Matching Function	Own Calculations, see section 4.3	0.576
$\beta$	Workers' Bargaining Weight	Mortensen and Pissarides (1999)	0.5

Parameters based on the IBGE National Accounts employ simple averages between 2000 and 2005. WIOD stands for "World Input-Output Database."

## 4.2 Estimation Procedure

We take the parameters described in Table 6 as given and estimate the remaining parameters using an Indirect Inference procedure. In this step, we estimate 27 parameters using 74 data moments and auxiliary model coefficients, ensuring that all equilibrium conditions listed in section 3.7 are met throughout the procedure. The estimation algorithm is described in detail in Supplementary Material I, but we highlight here a few important features of the procedure.

First, rewrite the value-added functions (8) and (34) as:

$$VA_k(z, \ell) = \Theta_k \Psi_k \left( \exp(\mathcal{I}_k^x(z, \ell) d_F) \right)^{\frac{\sigma_k}{\sigma_k - 1} \Lambda_k} \left( z \ell^{\delta_k} \right)^{\Lambda_k}, \quad (36)$$

<sup>29</sup>See Online Appendix K for details.

where  $\Psi_k \equiv (P_k^m)^{-(1-\delta_k)\Lambda_k} (\exp(d_{H,k}))^{\frac{\sigma_k}{\sigma_k-1}\Lambda_k}$  for  $k = C, S$ .<sup>30</sup> In addition, define  $\vartheta_{J_u} \equiv b_0 \times P_Y + \frac{1}{1+r} J^u$  as the expected discounted present value of unemployment. The estimation procedure treats the endogenous equilibrium objects  $\Psi_C$ ,  $\Psi_S$ , and  $\vartheta_{J_u}$  as “parameters” to be estimated.<sup>31</sup> Given a guess of these objects and of the remaining structural parameters, we are able to solve, for the mass of entrants, the mass of active firms, the firm-level policy functions, the steady-state distribution of states, and the unemployment rate. This solution is achieved by ensuring that the guesses and structural parameters are consistent with the equilibrium conditions.

To be precise, we minimize the loss function  $\mathcal{L}$  given by:  $\mathcal{L}(\Omega) = \sum_i W_i |m_i^{Model}(\Omega) - m_i^{Data}|$ , where  $\Omega$  denotes the set of parameters to be estimated,  $m_i^{Data}$  denotes the values of moments or auxiliary model coefficients measured in the data,  $m_i^{Model}(\Omega)$  denotes the values of moments or auxiliary model coefficients generated by the model when the set of parameters is given by  $\Omega$ , and  $W_i$  weights the importance of moment  $i$  in the loss function.  $\Omega$  includes the endogenous objects  $\Psi_C$ ,  $\Psi_S$ , and  $\vartheta_{J_u}$ , but excludes the structural parameters  $b_0$  and  $D_F^*$ , which are obtained after the minimization of  $\mathcal{L}$  is complete. It also excludes parameters  $\delta_k$  ( $k = C, S$ ), which are directly determined as functions of the elasticities of substitutions  $\sigma_k$  and the share of gross revenues devoted to intermediate goods payments using:  $\delta_k = 1 - \frac{\sigma_k}{\sigma_k-1} \left( \frac{Total\ Expenditures\ with\ Intermediates_k}{Total\ Gross\ Revenues_k} \right)^{Data}$ , where the term in parentheses is computed from the Brazilian National Accounts compiled by IBGE. Finally,  $\Omega$  also omits the foreign demand shifter  $d_F$ . Given a guess of  $\sigma_C$ ,  $d_F$  can be directly recovered by means of equation (33) and data on the average share of exporters’ gross revenues that is actually exported (which are obtained from PIA and SECEX).

Once the minimization of  $\mathcal{L}$  over  $\Omega$  is concluded, the flow utility of unemployment is recovered with:

$$b_0 \times P_Y = \vartheta_{J_u} - \frac{1}{1+r} J^u,$$

where  $J^u$  can be computed by means of equation (I.3) in the Online Appendix, and  $P_Y$  is the price of the final good (which is also a by-product of the estimation procedure). In turn,  $D_F^*$  is recovered as:

$$D_F^* = \frac{(\exp(\sigma_C \times d_F) - 1) (P_C^m)^{(1-\delta_C)(\sigma_C-1)} \Psi_C^{\frac{\sigma_C-1}{\Lambda_C}}}{\bar{\epsilon}^{\sigma_C} \tau_C^{1-\sigma_C}},$$

where  $\bar{\epsilon}$  is the exchange rate value that balances trade and  $P_C^m$  is the price of one unit of sector  $C$ ’s intermediate bundle.

We emphasize that the only differences between formal and informal firms are that the former are subject to regulations and taxes, and the latter face an expected cost of informality  $K^{inf}(z, \ell')$ . Otherwise, they have access to the same production and hiring technologies, and are subject to the

---

<sup>30</sup>  $\Theta_k \equiv \left( \frac{1}{(1-\delta_k)\Lambda_k} \right) \left( \frac{(1-\delta_k)(\sigma_k-1)}{\sigma_k} \right)^{\frac{\sigma_k-1}{\sigma_k-1}\Lambda_k}$  and  $\Lambda_k \equiv \frac{\sigma_k-1}{\sigma_k-(1-\delta_k)(\sigma_k-1)}$  are sector-specific constants;  $P_k^m$  is the price of one unit of sector  $k$ ’s intermediate bundle; and  $d_{H,k} \equiv \log \left( \left( \frac{X_k}{P_k^{1-\sigma_k}} \right)^{\frac{1}{\sigma_k}} \right)$  are domestic demand shifters. Finally, we impose  $\mathcal{I}_S^x(z, \ell) = 0$ —no firm in the service sector exports—to simplify the notation.

<sup>31</sup> In counterfactual simulations, these are solved out under the new environment. Estimation relies on the assumption that the data come from an equilibrium allocation.



same exogenous exit rates. However, these parameter restrictions do not significantly affect the final fit of the model.

### 4.3 Identification

To understand which moments in the data identify the parameters we estimate, first rewrite the hiring functions as:  $H_k(\ell, \ell') = \tilde{h}_k \left( \frac{\ell' - \ell}{\ell^{\gamma_{k2}}} \right)^{\gamma_{k1}}$ , where  $\tilde{h}_k = (\mu^v)^{-\gamma_{k1}} \left( \frac{h_k}{\gamma_{k1}} \right)$ , for  $k = C, S$ . Our estimation procedure treats  $\tilde{h}_k$  as a “parameter” to be estimated. This term is identified based on the average level of turnover rates at the firm level as well as the unemployment rate, given that the two are closely related.<sup>32</sup> As we discussed in section 3.3, the auxiliary (linear) model relating firm-level turnover rates to log employment and an export indicator gives information on the scale economies  $\gamma_{k2}$ . The auxiliary (linear) model relating log wages to log employment and an export indicator gives information on the convexity of hiring costs  $\gamma_{k1}$ , as it relates to wage dispersion across firms with different characteristics.

Note that  $\tilde{h}_k$  is a combination of a structural parameter  $h_k$  with an endogenous term  $\mu^v$ . Following the estimation of  $\tilde{h}_k$  and  $\gamma_{k1}$ , we recover  $h_k$  by imposing that  $\mu^v = 0.5$  in the equilibrium that we estimate.<sup>33</sup> In turn, we set  $\phi$  so that we perfectly match the yearly transition rate from unemployment to employment in the data.<sup>34</sup>

The model needs the fixed costs of production  $\bar{c}_k$  to match how the probability of firm-level exit rates declines with size. Exogenous destruction rates  $\alpha_k$  are needed to match the average exit rates. Matching the relationship between firm-level revenues and firm size gives information on  $\sigma_k$ .<sup>35</sup> We pin down the AR(1) process for productivity by targeting two dynamic moments: the serial correlation in firm-level employment, and the serial correlation in firm-level revenues. The cross-sectional dispersions in firm-level employment size and revenues are also informative about the variance of shocks  $\sigma_k^z$ . The share of  $C$  sector firms that export pins down the fixed cost of exporting  $f_x$ .

Finally, we identify the cost-of-informality function  $\tilde{a}_k \exp \left\{ \tilde{b}_k (\ell - 1) \right\}$  by matching the firm-size distribution in the informal sector, the share of employment in the informal sector, and the fraction of informal firms conditional on employment size. All the moments and auxiliary models

<sup>32</sup>In our model, worker separations are followed by a period of unemployment. This mechanically ties turnover rates to unemployment rates.

<sup>33</sup>Simply put, we can identify the level of hiring costs  $\tilde{h}_k$  based on data on firm-level worker turnover. However, we cannot separate  $h_k$  from  $\mu^v$ . We impose  $\mu^v = 0.5$  to be able to recover  $h_k$ . Other assumptions could have been made. Similarly to how we treat  $\Psi_C$ ,  $\Psi_S$ , and  $\vartheta_{J_u}$ ,  $\mu^v$  is solved out in our subsequent counterfactual exercises.

<sup>34</sup>Equation (24) implies that transition rates from unemployment to employment (in any sector), in the model, are given by:  $\sum_{k,j} \mu_{kj}^e = \left( \frac{\phi}{(\mu^v)^\xi} \right)^{\frac{1}{1-\xi}}$ . We choose  $\phi$  so that our model perfectly matches the transition from unemployment to employment in the data,  $Transition_{Data}^{U \rightarrow E}$ , conditional on  $\mu^v = 0.5$  and  $\xi = 0.5$ . That is,  $\phi = \frac{(\mu^v)^\xi}{(Transition_{Data}^{U \rightarrow E})^{\xi-1}} = 0.576$ , as shown in Table 6.

<sup>35</sup>In a standard Melitz model with zero fixed costs of production, revenue per worker pins down the elasticity of substitution. Our model is significantly more complex, but the intuition that the relationship between revenue and optimal size depends on the elasticity of substitution remains. Rows 68 to 70 of Table Q.4 and Rows 71 and 72 of Table Q.6 in Online Appendix Q show that the relationship between log revenue and log employment size in either sector is sensitive to the sector-specific elasticity of substitution, as we argue here.



used in the estimation procedure (as well as the datasets we have used to compute each of them) are listed in Appendix M.

We conclude by highlighting that, apart from these heuristic arguments for identification, Online Appendix Q includes Figures Q.1 to Q.4, illustrating the local behavior of the loss function around each parameter estimate. In addition, Tables Q.1 to Q.6 in the same Online Appendix compute the elasticity of each moment targeted in the estimation with respect to each element of  $\Omega$ . These exercises corroborate that each parameter maps to each of the aforementioned moments.

## 4.4 Estimation Results

### 4.4.1 Estimates

Table 7 presents our parameter estimates. Informality costs as a fraction of revenues,  $p_{ki}(\ell)$ , differ across sectors. In the  $C$  sector, the informality costs start at a relatively low value but rapidly increase with size. On the other hand, the informality penalty starts at a larger value in the  $S$  sector, but increases with size at a slower pace. The different convexities of the informality penalty function in the  $C$  and  $S$  sectors are intuitive, as one would expect large manufacturing firms to face increasing hurdles to remaining invisible to the government.

Table 7: Parameter Estimates

Parameter	Description	$k = C$	$k = S$
$\tilde{a}_k$	Cost of Informality, Intercept	0.183	0.379
$\tilde{b}_k$	Cost of Informality, Convexity	0.090	0.011
$h_k$	Hiring Cost, Level	471.3	2,959.4
$\gamma_k^1$	Hiring Cost, Convexity	2.007	5.745
$\gamma_k^2$	Hiring Cost, Scale Economies	0.109	0.147
$\sigma_k$	Elasticity of Substitution	5.256	3.063
$\rho_k$	Productivity AR(1) Process, Persistence Coeff.	0.979	0.977
$\sigma_k^z$	Productivity AR(1) Process, Std. Dev. of Shock	0.197	0.340
$\alpha_k$	Exogenous Exit Probability	0.073	0.082
$\bar{c}_k$	Fixed Cost of Operation	174.404	29.383
$\delta_k$	Labor Share in Production	0.264	0.525
$c_k^e$	Entry Cost	5,004.2	3,960.6
$f_x$	Fixed Cost of Exporting	56,900.9	
$b_0$	Utility Flow of Unemployment	-0.764	
$(D_F^*)^{\frac{1}{\sigma_C}}$	Foreign Demand Shifter	1,169.5	

Notes: To aid interpretation,  $b_0$  is expressed as a fraction of real income per capita according to the model.

Our hiring function estimates display convexity ( $\gamma_C^1 = 2.0$ ,  $\gamma_S^1 = 5.7$ ) and scale economies ( $\gamma_C^2 = 0.11$ ,  $\gamma_S^2 = 0.15$ ) in both sectors. These results are in the same ballpark as the recent estimates from Coşar et al. (2016), who use data from Colombia. Similarly, our estimates of the elasticities of substitution in the  $C$  and  $S$  sectors, given by  $\sigma_C = 5.3$ , and  $\sigma_S = 3.1$ , respectively, are within the range in earlier papers (see, for example, Coşar et al., 2016; De Loecker and Warzynski, 2012; De Loecker et al., 2016; Broda and Weinstein, 2006). In particular, our estimates are consistent with those of Gervais and Jensen (2019) who find that services have elasticities of substitution

approximately one-quarter smaller than those of manufacturing.

Finally, we estimate a negative value for  $b_0$ , implying a significant disutility flow of unemployment. To aid its interpretation, Table 7 reports that  $b_0$  amounts to approximately three-quarters of real income per capita. This is equivalent to 2.3 times the real value of the 2003 minimum wage ( $\underline{w}/P_Y$ ). As a benchmark, also in the context of a search model, Meghir et al. (2015) estimate a negative value for  $b_0 \times P_Y$  for men in the São Paulo metropolitan region equal to 4.3 times the 2008 minimum wage.<sup>36</sup> As the search literature has shown, a negative value of unemployment is necessary to generate the magnitudes of wage dispersion typically found in the data (see, for example, Hornstein et al., 2011).

#### 4.4.2 Model Fit

Tables M.1 through M.7 in Online Appendix M evaluate how our model-generated moments and auxiliary model coefficients compare to those obtained in the data. Overall, our model with 27 structural parameters fits our 74 target data moments or auxiliary model coefficients well. Importantly, Facts 1 through 5 highlighted in section 2 are all well matched. Specifically, Table M.1 shows that employment shares and transition rates from unemployment are on target, so that our model comfortably replicates Fact 1. The sharp relationship between informal status and size (Fact 2) is also well reproduced by our model, as shown in the lower panel of Table M.7. Tables M.5 and M.7 show that average wages in the formal sector indeed exceed those in the informal sector, mirroring Fact 4. Fact 5, relating firm-level turnover to size and export status, is also well replicated as Table M.2 shows. Also noteworthy is the ability of the model to match how turnover rates relate to employment and export status conditional on both expansions and contractions. We do not directly target Fact 3, but we verify that the model implies that less productive firms tend to sort into informal status. Finally, Table M.5 shows that the model recreates the strong wage size and exporter premia found in the data, in both sectors. These are important moments to replicate, as they give us confidence in the wage inequality counterfactuals we conduct in section 5.4.

Meghir et al. (2015) emphasize an important feature of the data on formal and informal firms, namely that the distributions of worker productivities across these two types of firms overlap. There are two reasons why our model generates this overlap. First, and more importantly, we do not allow formal firms to switch to informal status. This generates hysteresis: formal firms (which are more productive than informal firms at the entry stage) keep their status even if their productivity significantly declines over time, falling below the entry cutoff for formal firms. Second, selection into formality/informality depends on a two-dimensional state  $(z, \ell)$ . Therefore, for a fixed value of  $z$ , we can have selection into informality for some levels of  $\ell$  and into formality for other values of  $\ell$ , also contributing to the overlap in the two productivity distributions. Online Appendix O shows the overlap between these distributions within manufacturing and services.

Having discussed the successes of the model, we now turn to the moments that are not as

---

<sup>36</sup>Table 5 in their paper estimates a (monthly) flow value of unemployment of -1,308 for men in São Paulo. The average monthly minimum wage during the period that they consider was of R\$ 300.

well matched, and discuss reasons for some of these mismatches. Although the model is able to replicate very well the dependence between firm-level turnover rates, firm-level employment, and export status, the *average* turnover rate is not as well replicated. There is a tension between matching the *level* of turnover rates and matching the unemployment rate.

Our model also tends to underestimate the dispersion of firm size in the formal sector for both sectors  $C$  and  $S$  (see Table M.3). We hypothesize two reasons for this discrepancy. First, the data source that we employ has a thicker-than-usual left tail of the firm-size distribution.<sup>37</sup> Second, the assumed normal distribution of productivity shocks naturally makes it harder to match the right tail of the size distribution. Although modeling firm-level dynamics according to equation (6) is standard in the literature (see Hopenhayn and Rogerson, 1993; Alessandria and Choi, 2014; Coşar et al., 2016, among many others), we would need a more flexible specification for the productivity process to better match the firm-size distribution in Brazil.<sup>38</sup> Relatedly, the dispersion in revenues in Table M.6 is also underestimated.

Finally, we overestimate average productivity/revenues among informal  $C$  sector firms. In practice, there is a tension between matching this moment and the level of wages in the informal  $C$  sector, given that we currently underestimate the latter. Nonetheless, our model-generated moments are consistent with the fact that informal firms are, on average, substantially smaller and less productive than formal firms.

## 5 Results

### 5.1 The Impact of Trade Costs on Real Income and Informality

#### 5.1.1 Small Changes in Trade Costs

We start by analyzing the behavior of the model in response to a small shock in iceberg trade costs  $\tau_c$ . Specifically, we compare the outcomes of our model with  $\tau_c = 2.3$  to those of the benchmark model under  $\tau_c = 2.4$ —holding all the remaining parameters constant at their calibrated values. This small shock allows us to decompose the total effect of a reduction in trade costs on real income into a “mechanical effect” and a “reallocation effect,” in the spirit of the decomposition of Baqaee and Farhi (2020).

The mechanical effect measures the real income benefits from lower trade costs or other technology changes that materialize solely through the lower cost of imported varieties and feedback effects into production costs through intersectoral linkages. Because we hold fixed initial labor allocations, wages, aggregate income, and the total demand that firms face, this effect reflects only the mechanical (or immediate) reduction in price indices implied by lower trade costs. On the other hand, the reallocation effect holds trade costs and technology parameters constant. This term reflects both

<sup>37</sup>The 20<sup>th</sup> and 40<sup>th</sup> percentiles of firm size in the formal  $C$  sector are given by, respectively, 2 and 4. In the formal  $S$  sector, these are 1 and 2.

<sup>38</sup>Although standard in the literature, imposing log-normality can have important quantitative implications for the gains from integration, as discussed in Adão et al. (2020).

general equilibrium effects and the consequences of reallocating resources across uses that are more or less distorted. If the underlying trade cost or technology shock is small enough, we can employ the following first-order approximation, which is the point of departure in the analysis of [Baqae and Farhi \(2020\)](#):

$$\Delta \text{Real Income} \approx \Delta \text{Mechanical Effect} + \Delta \text{Reallocation Effect}. \quad (37)$$

Once we feed the change in trade cost from  $\tau_c = 2.4$  to  $\tau_c = 2.3$  into the model, we can compute the resulting change in real income ( $\Delta \text{Real Income}$ ) that this shock entails and the mechanical effect associated with it ( $\Delta \text{Mechanical Effect}$ ). We describe the details behind the computation of the mechanical effect in Appendix J. The reallocation effect can therefore be backed out as a residual using equation (37).<sup>39</sup>

Table 8 displays the results of this decomposition for the small shock that we consider. The cells present percentage deviations relative to the benchmark equilibrium with  $\tau_c = 2.4$ . Reducing trade costs from  $\tau_c = 2.4$  to  $\tau_c = 2.3$  leads to a real income gain of approximately 1.3%. Strikingly, the reallocation effect is nearly five times greater than the mechanical effect (1.07% versus 0.23%), accounting for 82% of the real income gains resulting from increased globalization.

Table 8: Impacts of a Small Trade Cost Shock,  $\tau_c$  declines from 2.4 to 2.3

1. $\Delta \text{Real Income}$	1.296	
2. $\Delta \text{Mechanical Effect}$	0.228	
3. $\Delta \text{Reallocation Effect}$	1.067	
4. $\Delta \text{Unemployment}$	1.940	
	<i>C Sector</i>	<i>S Sector</i>
5. $\Delta \text{Share Informal Employment}_k$	-4.400	-1.315
6. $\Delta TFP_k$	1.637	1.502
7. $\Delta \text{Var}(\log(\partial R_k(z, \ell)/\partial \ell))$	-1.764	-1.898

All numbers are expressed in percentage deviations relative to the Benchmark with  $\tau_c = 2.4$ .  $TFP_k$  is the aggregate total factor productivity in sector  $k$ , defined as the average of firm-level productivities  $z$  weighted by employment.

As discussed at length in [Baqae and Farhi \(2020\)](#), in an economy without distortions, the reallocation effect would be negligible for small shocks. However, as we previously highlighted, our setting has many distortions. Crucially, because taxes and regulations are not perfectly enforced by the government, firms are not equally exposed to these domestic distortions. Consequently, smaller and less efficient informal firms experience lower distortions, leading them to produce above the socially optimal level. In contrast, larger and more productive formal firms face greater distortions, prompting them to produce below the socially optimal level.

A key consequence of a reduction in trade costs is that it leads to a reduction in the share of

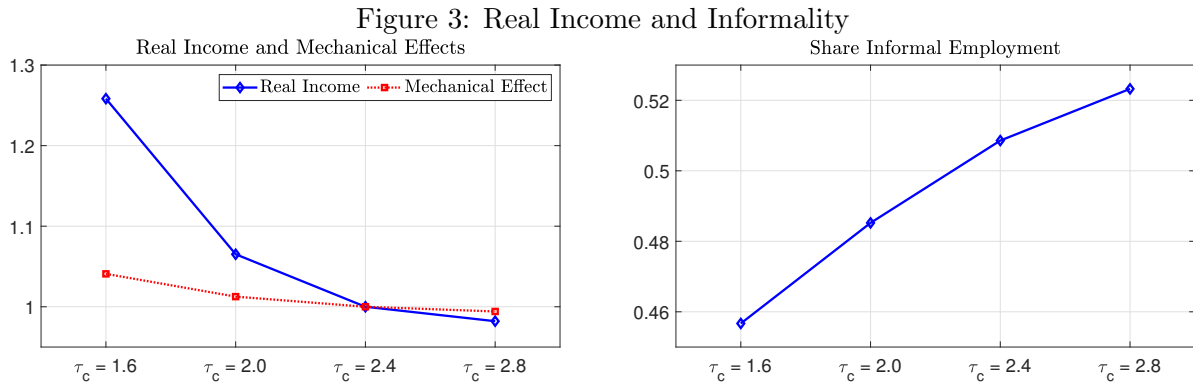
<sup>39</sup>[Baqae and Farhi \(2020\)](#) develop analytical expressions for both the mechanical effect and the reallocation effect. Their expressions directly connect observed data to these effects. Unfortunately, this is not possible in our setting, which is substantially more complex to theirs. Important departures from their assumptions include: (a) heterogeneous wages across firms for homogeneous workers; (b) search frictions and unemployment; (c) firm dynamics subject to adjustment costs; (d) minimum wages and payroll taxes; (e) endogenous entry and exit of firms.

informal employment in both  $C$  and  $S$  sectors, as is illustrated in row 5 of Table 8. Concurrently, we observe a reallocation of labor from less distorted, low-productivity informal firms to more distorted, high-productivity formal firms. This is highlighted in row 6 of Table 8 which shows the behavior of measured aggregate TFP in both sectors.<sup>40</sup> Finally, row 7 shows a reduction in the dispersion of log marginal revenue product of labor in both sectors—suggesting a more efficient allocation of resources following increases in globalization, further boosting the reallocation effect (Hsieh and Klenow, 2009).

### 5.1.2 Larger Changes in Trade Costs

The impacts that we document in Table 8 focus on small changes in trade costs. Armed with this small shock, we are able to leverage the decomposition in equation (37) to back out the reallocation effect and compare it with the mechanical effect of an increase in trade openness. The procedure to compute the mechanical effect that we outline in Online Appendix J is still valid for large shocks. However, with large shocks, the first-order approximation in equation (37) is no longer accurate and we are unable to back out the reallocation effect.

Nevertheless, we simulate the impacts of larger shocks to trade costs as these are the most relevant in practice. This section focuses on reductions in trade costs from  $\tau_c = 2.4$  to  $\tau_c = 2.0$  and  $\tau_c = 1.6$  as well as increases to  $\tau_c = 2.8$ . To understand the magnitude of the implications of these trade cost changes, note that our benchmark equilibrium with  $\tau_c = 2.4$  implies an imports-to-GDP ratio of 6.6%. Reducing  $\tau_c$  to 2.0 leads to an imports-to-GDP ratio of 15%, and further reducing it to 1.6 implies an openness ratio of 30%, similar to the ratios currently observed in the United Kingdom, South Africa, and Colombia (World Bank, 2022).



Note: Real Income and the Mechanical Effect on the left panel are displayed relative to their benchmark values under  $\tau_c = 2.4$ .

The left panel of Figure 3 shows that reducing trade costs from  $\tau_c = 2.4$  to  $\tau_c = 2.0$  leads to a total real income gain of 6.5%, whereas the mechanical effect is 1.3%. Next, reducing trade costs from  $\tau_c = 2.4$  to  $\tau_c = 1.6$  leads to a total real income gain of 25.8%, whereas the mechanical

<sup>40</sup>By “measured aggregate TFP,” we refer to the weighted average of firm-level productivities  $z$ , where the weights are given by firm-level employment. This is a commonly used empirical measure of aggregate TFP in the literature (Olley and Pakes, 1996; Pavcnik, 2002). However, this measure is not necessarily informative about welfare, as discussed in Berthou et al. (2021).

effect is of 4.1%. In both cases, the total real income gains greatly exceed the mechanical effect, corroborating our previous conclusion that the reallocation effect is quantitatively important in our context of widespread informality. As  $\tau_c$  declines, similarly to the impacts documented in Table 8, the informal sector contracts. There is a reallocation of labor from initially less distorted but less productive informal firms to initially more distorted but more productive formal firms. Additionally, the dispersion in the log marginal revenue product of labor declines.

We conclude this subsection by interpreting the magnitude of the impact of reductions in trade costs on the size of the informal sector. The right panel of Figure 3 shows that, although the impacts on the total number of informal workers are economically important, they are not enough to *significantly* curb informality. Specifically, as  $\tau_c$  declines from 2.4 to 1.6, the imports-to-GDP ratio increases from 6.6% to 30%, a substantial fourfold increase. However, the share of informal workers declines much more modestly, from 51% to 46%.

While the overall share of informality decreases by 5 percentage points, the fraction of informal employment within the manufacturing sector experiences a more substantial decline of 25%. This observation, coupled with our roundabout production structure and intersectoral linkages, helps to explain the significant reallocation effect that we uncover.

### 5.1.3 Trade, Informality and Unemployment: Mechanisms

We conclude this subsection by discussing the impacts on informality and unemployment, as shown in Table 8. It is important to note that the same mechanisms and forces come into play when we subject the model to larger shocks.

We begin by dissecting the effects of reducing trade costs on informality in sector  $C$ . When trade costs decrease, the foreign market size  $d_F$  increases, leading exporters to grow, become more profitable, and pay higher wages. In addition, the entry of new firms into the export market results in an increased share of firms paying higher wages. These impacts raise the value of search  $J^u$ , causing wages to rise across all firms in the economy. Consequently, these increased wages force the least productive informal firms out of the market. This mechanism tends to reduce informality in sector  $C$ , all else being equal.

Next, we have two opposing forces. The aforementioned higher wages also tend to push the least productive formal firms out of the market as their labor costs increase. *Ceteris paribus*, this would tend to increase informality. However, high-productivity informal firms anticipate that, at some point, they may grow and become exporters with higher probability. This increases the value of formality relative to informality, encouraging high-productivity informal firms to transition to the formal sector. This tends to decrease informality. Quantitatively, the combined effect of these forces is a higher productivity threshold for entry in the informal sector and a lower threshold for entry in the formal sector, unambiguously resulting in a reduction in the size of the informal sector in sector  $C$ .

The mechanisms at play in sector  $S$  are similar to those in sector  $C$ , except that high-productivity informal sector firms cannot export. As a result, we observe in our simulation an increase in the

productivity threshold for entry in both the informal and formal sectors simultaneously. The impact of trade openness on informality in the  $S$  sector is therefore ambiguous. However, quantitatively, there is a net decline in the informal sector in the  $S$  sector as well.

Now, we turn to the unemployment impact of the reduction in trade costs. Table 8 shows that the unemployment rate increases by 1.9%. To understand this effect, remember that we have two opposing forces connecting globalization to unemployment: as trade costs decline there is (i) a reallocation of resources toward larger and more stable firms; and (ii) a reallocation of resources toward exporters whose revenues become more volatile given the larger foreign market size  $d_F$  and whose share in the economy increases through an extensive-margin adjustment. Depending on their relative forces, these mechanisms can lead to either an increase or a decline in the aggregate turnover rate. Quantitatively, the second force dominates, leading to higher turnover in the economy. Given search frictions, higher aggregate turnover is translated into higher unemployment rates. One additional mechanism linking globalization to unemployment operates through the aforementioned increased wages, which negatively affect the number of vacancies posted, especially at the bottom of the productivity distribution.

We emphasize that this increase in unemployment only partially offsets the real income gains from trade, which are still large. In results available upon request, we show that the impact of globalization on unemployment is also positive for the larger trade cost shocks that we simulate in section 5.1.2. As trade costs are reduced from  $\tau_c = 2.4$  to  $\tau_c = 1.6$ , the unemployment rate increases from 18.6% to 21%.<sup>41</sup>

## 5.2 The Impact of Trade Costs under Various Enforcement Scenarios

The previous section has shown that, in our context of high informality, trade openness leads to a large reallocation effect that greatly exceeds the mechanical effect. We now investigate the role of the informal sector in the substantial reallocation effect that we have uncovered. To do so, we gradually increase the costs that informal firms face, slowly reducing the share of firms and workers operating in the informal sector. We interpret these increases in the costs of informality as government enforcement gradually becoming stricter, with informality penalized more and more. Eventually, enforcement becomes perfect—so that the costs of operating informally are infinite—and no firm or worker operates informally. Throughout this process, all the remaining parameters of the model are held constant at their calibrated benchmark values in Table 7.

More precisely, we consider five scenarios of enforcement: the benchmark case that we estimate, “Benchmark”; the perfect-enforcement scenario with no informality, “No Informality”; and three scenarios in between these two extremes, “Stricter Enforcement 1, 2 and 3”. Recall that informality costs are given by  $K^{inf}(z, \ell) = p_{ki}(\ell) \times R_k(z, \ell)$ ,  $k \in \{C, S\}$ . Figure 4 illustrates the intermediate scenarios in enforcement  $p_{ki}(\ell)$  that we simulate. As we move from the “Benchmark” scenario to

---

<sup>41</sup>We have been focusing on the real income impacts of trade costs. However, unemployment brings disutility to workers, as we discussed in section 4.4. We note that the increase in unemployment and associated disutility is modest relative to the real income gains, so that the gains from trade net of the disutility of unemployment are still large and positive.



the “No Informality” case,  $p_{ki}$  becomes gradually steeper or shifts upward.

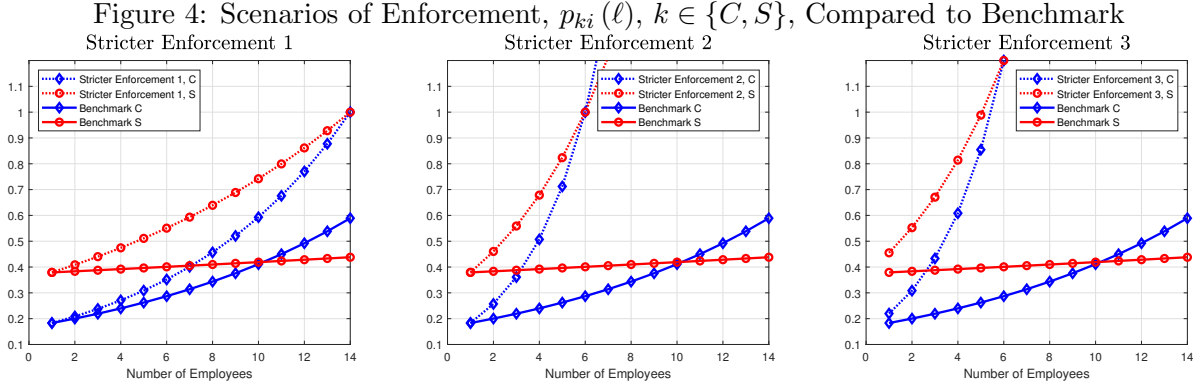


Table 9 shows the impact of gradually increasing informality costs in both sectors on the share of informal employment. As the first row highlights, the “Benchmark” economy features 51% of employment in the informal sector. This rate declines to approximately 36% under the scenario “Stricter Enforcement 1”, then down to 29% under “Stricter Enforcement 2”, 16% under “Stricter Enforcement 3”, and finally 0 when the government perfectly enforces taxes and regulations under “No Informality”. Rows 2 and 3 of Table 9 break these effects down across sectors  $C$  and  $S$ . It is noteworthy that, as we gradually increase enforcement, the unemployment rates also gradually increase from 18.6% under “Benchmark” up to 30% under “No Informality”, when no informality is allowed. The primary reason for the significant increase in unemployment between these two extreme scenarios is that the minimum wage becomes increasingly binding as firms find it more and more costly to evade it by choosing to operate in the informal sector. Finally, the last row of Table 9 shows that these five scenarios feature similar levels of openness, which implies that the mechanical effect of a reduction in trade costs is similar across scenarios.

Table 9: Effects of Increasing the Cost of Informality

	Bench.	Strict Enf. 1	Strict Enf. 2	Strict Enf. 3	No Inf.
1. Share Informal Emp.	0.509	0.358	0.288	0.157	0.000
2. Share Emp. $Ci$	0.076	0.062	0.042	0.029	0.000
3. Share Emp. $Si$	0.432	0.296	0.246	0.128	0.000
4. Unemployment Rate	0.186	0.207	0.221	0.253	0.303
5. Imports/GDP	0.066	0.069	0.073	0.076	0.084

Table 10 shows the impacts of a small reduction in trade costs from  $\tau_c = 2.4$  to  $\tau_c = 2.3$  within each of the various enforcement scenarios. We first note that the real income gains under the “Benchmark” economy, with a large informal sector, are larger than the gains in all the remaining scenarios when the informal sector is repressed. Furthermore, note that the mechanical effect is similar across scenarios, meaning that the differences in real income gains from globalization are driven by differences in the magnitude of the reallocation effect. Although the reallocation effect is not monotonically decreasing with the level of enforcement, it is uniformly lower in all economies



with a repressed informal sector.<sup>42</sup> These lower reallocation effects are driven by a less pronounced extent of reallocation from low-productivity, less distorted informal firms to high-productivity, more distorted formal firms. This is well illustrated in rows 4 and 5 of Table 10 which show that the impact of trade openness on aggregate TFP declines with the level of enforcement. Similarly, the impact of trade openness on the dispersion of the log marginal revenue product of labor tends to be more negative in the “Benchmark” economy. The main conclusion is that an increase in globalization leads to a reallocation of labor and output from less distorted, initially overproducing informal firms to more distorted, initially underproducing formal firms and that this boosts the reallocation effect. In an economy with higher levels of enforcement, trade openness has less room to accomplish this reallocation and therefore the magnitude of the reallocation effect tends to be smaller.

Table 10: Impacts of a Small Trade Shock,  $\tau_c$  declines from 2.4 to 2.3

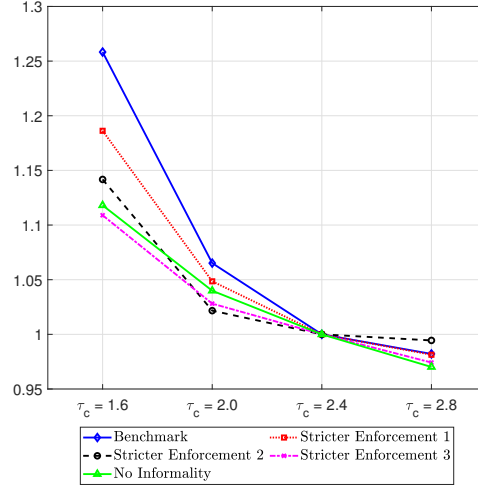
	Bench.	Strict Enf. 1	Strict Enf. 2	Strict Enf. 3	No Inf.
1. $100 \times \Delta Real\ Income$	1.296	0.794	0.554	0.634	0.872
2. $100 \times \Delta Mechanical\ Effect$	0.228	0.239	0.253	0.263	0.286
3. $100 \times \Delta Reallocation\ Effect$	1.067	0.555	0.301	0.372	0.585
4. $\Delta TFP_C$	1.637	1.367	1.016	0.901	0.623
5. $\Delta TFP_S$	1.502	0.149	0.044	0.048	-0.299
6. $\Delta Var(\log(\partial R_C(z, \ell)/\partial \ell))$	-1.764	-1.832	0.219	-0.133	3.274
7. $\Delta Var(\log(\partial R_S(z, \ell)/\partial \ell))$	-1.898	-0.087	-0.096	-0.056	0.353

All numbers are expressed in percentage deviations relative to  $\tau_c = 2.4$ , and conditional on the specific enforcement scenario.  $TFP_k$  is the aggregate total factor productivity in sector  $k$ , defined as the average of firm-level productivities  $z$  weighted by employment.

Figure 5 moves away from small  $\tau_c$  shocks to simulate the larger shocks considered in section 5.1.2. Consistent with Table 10, we confirm that the real income gains from trade are considerably larger in the “Benchmark” economy with a high degree of informality. As trade costs decrease from  $\tau_c = 2.4$  to  $\tau_c = 1.6$ , the “Benchmark” economy witnesses real income gains of 25.8%. An economy with a slight repression of the informal sector, “Stricter Enforcement 1”, experiences gains of 18.6%. Finally, the “No Informality” scenario shows gains of 11.8%, which is less than half of those accrued in the “Benchmark” economy. The main conclusion from this section is that the gains from trade liberalization tend to be larger in economies with larger informal sectors. The main reason is that trade reallocates resources from initially less distorted but overproducing firms to initially more distorted underproducing firms. In economies where the government can better enforce compliance with taxes and regulations, there is less room to gain from such reallocation.

<sup>42</sup>The substantially larger reallocation effect under “No Informality” than under “Stricter Enforcement 3” (0.585 vs. 0.372) is explained by differences in the behavior of the unemployment rate in these two scenarios. The unemployment rate is approximately stable under “Stricter Enforcement 3”, but it sharply decreases under “No Informality”. The reason is that, as we explained earlier, minimum wages are severely binding under the “No Informality” scenario. Once trade costs are reduced, firms across the board offer higher wages in equilibrium, so that the minimum wage becomes less binding and exerts less pressure on unemployment. In additional results available upon request, we show that the impact of the trade cost shock on real income per *employed* worker approximately declines monotonically with the degree of enforcement.

Figure 5: Effect of Trade on Real Income Relative to  $\tau_c = 2.4$ ,  
Various Enforcement Scenarios



### 5.3 Comparison with Earlier Reduced-Form Work: The Effects of Negative Labor Demand Shocks

Our paper also relates to a growing body of empirical work that highlights the responsiveness of the informal sector to trade shocks in developing countries (Goldberg and Pavcnik, 2003; McCaig and Pavcnik, 2018; Dix-Carneiro and Kovak, 2019). More broadly, there is an extensive literature exploring the local labor market impacts of foreign competition. However, a significant portion of these studies relies on differences-in-differences specifications, which are not conducive to assessing the aggregate effects of trade (see, e.g., Adão et al., 2023; Dix-Carneiro and Kovak, 2023). These analyses are focused on a distinct question, namely, how (trade-induced) sector- or region-specific *labor demand shocks* influence *relative* sector- or region-specific informality or other labor market outcomes.

It is therefore important to underscore that the trade-induced local impacts documented in, for example, Dix-Carneiro and Kovak (2019) do *not* measure the aggregate impacts of trade on informality. Their specification simply compares labor market outcomes of a region facing a smaller (trade-induced) negative labor demand shock to those of a region facing a larger (trade-induced) negative labor demand shock. Given that common impacts of trade are absorbed into the intercept, this specification isolates only the impacts of negative labor demand shocks. From our model’s standpoint, this involves comparing the “Benchmark” economy with an economy experiencing a negative labor demand shock. This analysis is important as it directly addresses the empirical work of Dix-Carneiro and Kovak (2019) and Ponczek and Ulyssea (2022), enabling us to contrast the predictions of our model with their empirical results. We perform this comparison by simulating negative aggregate productivity shocks where we uniformly shift the support of the whole productivity process to the left. Here, we focus our analysis on a small shock, shifting the productivity support to the left by 0.5%, enabling us to once again explore the first-order approximation in

equation (37).<sup>43</sup>

Table 11: Small shock, Negative Productivity Shock of 0.5 pct

	Bench.	Strict Enf. 1	Strict Enf. 2	Strict Enf. 3	No Inf.
1. $100 \times \Delta \text{Share Informal Employment}$	5.106	4.106	1.407	3.141	—
2. $100 \times \Delta \text{Unemployment}$	-1.413	-0.365	0.929	1.767	3.530
3. $100 \times \Delta \text{Real Income}$	-4.688	-2.457	-1.509	-1.814	-2.666
4. $100 \times \Delta \text{Mechanical Effect}$	-0.308	-0.308	-0.306	-0.306	-0.304
5. $100 \times \Delta \text{Reallocation Effect}$	-4.379	-2.149	-1.202	-1.508	-2.362
6. $100 \times \Delta \text{Agg TFP}_C$	-1.706	-1.257	-0.806	-0.062	0.593
7. $100 \times \Delta \text{Agg TFP}_S$	-4.210	-2.027	-0.354	-0.018	1.612
8. $100 \times \Delta \text{Var}(\log(\partial R_C(z, \ell)/\partial \ell))$	1.021	1.303	1.887	-0.520	-5.796
9. $100 \times \Delta \text{Var}(\log(\partial R_S(z, \ell)/\partial \ell))$	4.315	2.067	0.087	0.145	-3.355

All numbers are expressed in percentage deviations relative to their respective initial equilibrium (column by column).  $TFP_k$  is the aggregate total factor productivity in sector  $k$ , defined as the average of firm-level productivities  $z$  weighted by employment.

Rows 1 and 2 of Table 11 show that an economy-wide negative labor demand shock has a relatively modest (and actually negative) effect on unemployment in the “Benchmark” economy. However, it has a large positive impact on the share of informal employment, increasing it by 5%. These results parallel those of [Dix-Carneiro and Kovak \(2019\)](#) who show that, in the context of the Brazilian trade liberalization, regions facing more negative labor demand shocks experience no discernible responses of unemployment in the long run but do experience substantial relative increases in the share of labor informally employed. This result prompted the authors to propose a hypothesis: that the informal sector might have served as a fallback for displaced workers.

This hypothesis is further explored by [Ponczek and Ulyssea \(2022\)](#), who revisit the findings of [Dix-Carneiro and Kovak \(2019\)](#). They find that the effects of trade-induced declines in labor demand on unemployment were more pronounced in regions where labor markets were closely monitored by the government (i.e., where enforcement of regulations was stricter), while the impacts on informality were less significant. These empirical results align with the findings presented in Table 11, indicating that in economies where the informal sector is more tightly monitored, the impact of the negative labor demand shock on informality tends to be smaller, while the effect on unemployment tends to be larger (see rows 1 and 2). Taken as a whole, the results from the model corroborate the informal sector’s role as an “unemployment buffer,” consistent with existing empirical evidence. This lends further credibility to our quantitative analyses in the previous sections.

Interestingly, the impact of the negative productivity shock on real income is more negative in the “Benchmark” economy than in the other economies where the informal sector is repressed—see row 3. The primary driver behind this result is the reallocation effect, which, in the “Benchmark” case, is over 10 times larger than the mechanical effect. Rows 6 to 7 indicate a negative impact of the shock on aggregate TFP. This effect is significantly larger than the shock itself and is driven

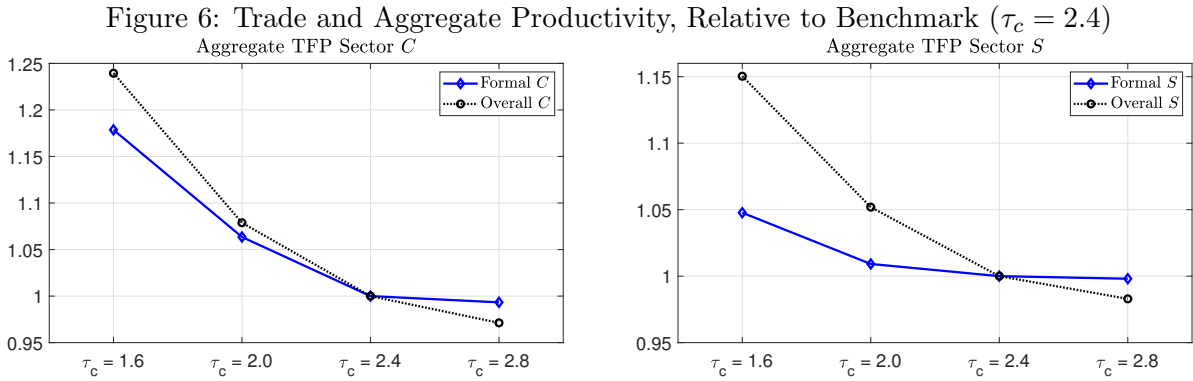
<sup>43</sup>We also conduct analyses with larger shocks of 1.5% and 3%. These results are available upon request. The insights gained from the small shock analysis remain applicable to the larger shocks.

by a robust reallocation of labor to less distorted but less productive firms in the informal sector, thereby amplifying misallocation in the economy.

Rows 8 and 9 also reveal an increase in the dispersion of log marginal revenue product of labor in both sectors, further contributing to the misallocation of resources. The extent of this misallocation increase is mitigated in economies where the informal sector is repressed. The conclusion is that the informal sector indeed serves as a fallback for displaced workers, but this comes at the expense of an expansion of the informal sector, amplifying misallocation in the economy and further reducing welfare.

#### 5.4 Reassessing the Impact of Trade on Aggregate TFP and Wage Inequality

We conclude this section by emphasizing that our work also informs the interpretation of a range of empirical results focused on the formal sector when the economy has a significant informal sector—typically the case in many developing countries (particularly in Latin America). We begin with Figure 6, which illustrates that the conventionally measured TFP gains, based on data solely from the formal sector, fall short of capturing the overall TFP gains including the informal sector.<sup>44</sup> Studies that focus solely on the formal sector miss that reductions in trade costs force unproductive informal firms out of the market, amplifying the reallocation of labor from low- to high-productivity firms.



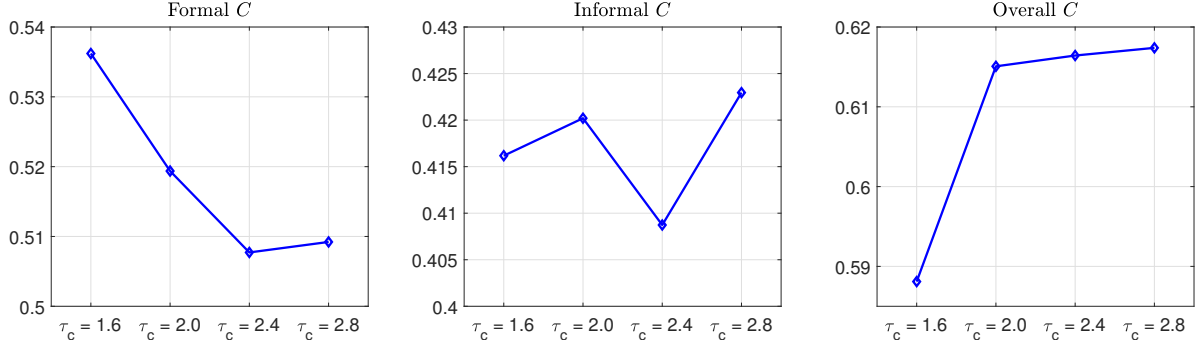
Next, we shift gears and revisit a standard finding in the international trade literature: namely, that trade liberalization increases wage inequality across firms in the manufacturing sector when we focus on formal-sector outcomes. Examples of this finding appear in [Helpman et al. \(2017\)](#), who focus on Brazil, and [Coşar et al. \(2016\)](#), who study the Colombian experience. Both papers focus on the effects of a reduction in trade costs on wage inequality within the formal manufacturing sector. They argue that in their respective contexts, the rise in trade openness has led to an increase in wage inequality across firms.

The left panel of Figure 7 shows that our model replicates these findings if we restrict attention to the evolution of wage inequality across firms. However, the right panel shows that reductions in trade costs lead to a decline in overall wage inequality across firms, once we take informal sector

<sup>44</sup>As before, empirically measured aggregate TFP is computed as the employment-weighted average of firm-level productivities  $z$ .

firms into account. This finding is explained by the fact that wage inequality across firms is reduced within the informal sector (middle panel) and that the gap between average wages in the formal and informal sectors also declines.<sup>45</sup>

Figure 7: Trade and Wage Inequality in Sector  $C$ , Standard Deviation of log Wages Across Workers



The results in this and previous sections underscore the importance of incorporating the informal sector into our models of trade. Not only does the magnitude of the gains from trade depend on the informal sector, but so do the interpretation and contextualization of a large range of empirical results for developing countries in which the researchers focused their attention on the formal manufacturing sector.

## 6 Conclusion

Understanding the implications of domestic distortions for the effects of international trade is an important area of research, particularly in developing countries where these distortions are expected to be significant (Atkin and Khandelwal, 2020; Atkin et al., 2022). In this paper, we investigate a set of distortions associated with the phenomenon of informality, namely burdensome labor market regulations and taxes that are imperfectly enforced by the government. This imperfect enforcement implies that some firms are subject to these regulations, while others are not. The first set of firms constitutes the formal sector, which, due to this imperfect enforcement, is more distorted than the informal sector.

We show that trade liberalization has large, positive impacts in this environment. However, only a small part of these beneficial effects can be attributed to a pure “mechanical effect.” Trade openness helps reallocate resources from the initially less distorted informal firms to the initially more distorted formal firms, leading to a sizable “reallocation effect.” Furthermore, we find that this effect is particularly pronounced in economies with a larger informal sector. Crucially, our analysis reveals that while the mechanical effect of trade remains relatively constant across scenarios, the reallocation effect diminishes as enforcement improves. In economies with stricter enforcement, the initial misallocation associated with the informal sector is smaller, thereby limiting the potential

<sup>45</sup>Wage inequality across all firms, encompassing both the formal and informal sectors, is similarly reduced in the  $S$  sector. Additionally, we observe reductions in inequality within both sectors and across all workers in lifetime worker values  $J^e(z, \ell)$ , which incorporates the utility flow of unemployment  $b_0$ .

for trade openness to facilitate the reallocation from less to more distorted firms. These results suggest that globalization is particularly valuable when the misallocation resulting from imperfect enforcement is largest.

In comparisons with earlier reduced-form work, we replicate the finding that informality can act as an employment buffer in the presence of negative labor demand shocks. However, our analysis reveals that informality does not act as an aggregate welfare buffer, as real income declines less in the presence of stricter enforcement despite the larger increase in unemployment. This is explained by the fact that negative economic shocks reallocate resources from more to less distorted sectors, exacerbating misallocation.

Importantly, our results underscore the importance of incorporating the informal sector in analyses of trade policies in developing countries, both quantitatively and qualitatively. We hope that our paper provides a first step in this direction by analyzing the microfoundations of the informality phenomenon, specifically the unequal exposure to domestic distortions and their interactions with trade.

We end the paper highlighting that the quantitative framework we present here is very flexible and amenable to studying the impacts of tax policies and labor market regulations in an environment with widespread informality, tax evasion and noncompliance. However, an important limitation of this paper lies in the fact that workers are assumed to be homogeneous in our model. It has been documented that the informal sector tends to be more unskilled labor intensive than the formal sector (Perry et al., 2007; La Porta and Shleifer, 2014), so policies affecting this sector can have important distributional consequences. We leave this extension for future work.

## References

- ADÃO, R., C. ARKOLAKIS, AND F. ESPOSITO (2023): “General Equilibrium Effects in Space: Theory and Measurement,” Unpublished Manuscript.
- ADÃO, R., C. ARKOLAKIS, AND S. GANAPATI (2020): “Aggregate Implications of Firm Heterogeneity: A Nonparametric Analysis of Monopolistic Competition Trade Models,” Unpublished Manuscript.
- ALESSANDRIA, G. AND H. CHOI (2014): “Establishment Heterogeneity, Exporter Dynamics, and the Effects of Trade Liberalization,” *Journal of International Economics*, 94, 207 – 223.
- ALVAREZ, J., F. BENGURIA, N. ENGBOM, AND C. MOSER (2018): “Firms and the decline in earnings inequality in Brazil,” *American Economic Journal: Macroeconomics*, 10, 149–89.
- ATKIN, D. AND D. DONALDSON (2022): “The Role of Trade in Economic Development,” in *Handbook of International Economics: International Trade*, ed. by G. Gopinath, E. Helpman, and K. Rogoff, Elsevier, vol. 5, 1–59.
- ATKIN, D., A. KHANDELWAL, L. BOUDREAU, R. DIX-CARNEIRO, I. MANELICI, P. MEDINA, B. MCCAIG, A. MORJARIA, L. PASCALI, B. RIJKERS, AND M. STARTZ (2022): “International Trade,” *VoxDevLit*, 4.
- ATKIN, D. AND A. K. KHANDELWAL (2020): “How distortions alter the impacts of international trade in developing countries,” *Annual Review of Economics*, 12, 213–238.
- BAI, Y., K. JIN, AND D. LU (2023): “Misallocation Under Trade Liberalization,” Working Paper.
- BANERJEE, A. V. AND E. DUFLO (2005): “Growth theory through the lens of development economics,” *Handbook of economic growth*, 1, 473–552.

- BAQAEE, D. R. AND E. FARHI (2020): “Productivity and Misallocation in General Equilibrium,” *The Quarterly Journal of Economics*, 135, 105–163.
- BERTHOUE, A., J. H. CHUNG, K. MANOVA, AND C. SANDOZ (2021): “Trade, Productivity and (Mis)allocation,” Working Paper.
- 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, 1339–1382.
- BRODA, C. AND D. E. WEINSTEIN (2006): “Globalization and the Gains From Variety,” *The Quarterly Journal of Economics*, 121, 541–585.
- CARD, D., A. R. CARDOSO, AND P. KLINE (2016): “Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women,” *The Quarterly journal of economics*, 131, 633–686.
- CARD, D., J. HEINING, AND P. KLINE (2013): “Workplace heterogeneity and the rise of West German wage inequality,” *The Quarterly journal of economics*, 128, 967–1015.
- CATÃO, L., C. PAGÉS, AND M. ROSALES (2009): “Financial Dependence, Formal Credit, and Informal Jobs: New Evidence from Brazilian Household Data,” IZA Discussion Paper No. 4609.
- COŞAR, A., N. GUNER, AND J. TYBOUT (2016): “Firm Dynamics, Job Turnover, and Wage Distributions in an Open Economy,” *The American Economic Review*, 106, 625–663.
- DE LOECKER, J., P. K. GOLDBERG, A. K. KHANDELOWAL, AND N. PAVCNİK (2016): “Prices, Markups, and Trade Reform,” *Econometrica*, 84, 445–510.
- DE LOECKER, J. AND F. WARZYŃSKI (2012): “Markups and Firm-Level Export Status,” *American Economic Review*, 102, 2437–71.
- DIX-CARNEIRO, R. (2014): “Trade Liberalization and Labor Market Dynamics,” *Econometrica*, 82, 825–885.
- DIX-CARNEIRO, R. AND B. K. KOVAK (2017): “Trade Liberalization and Regional Dynamics,” *American Economic Review*, 107, 2908–46.
- (2019): “Margins of Labor Market Adjustment to Trade,” *Journal of International Economics*, 117.
- (2023): “Globalization and Inequality in Latin America,” NBER Working Paper No. 31459.
- DOING BUSINESS (2007): “Paying Taxes - The Global Picture,” The World Bank: Washington, DC.
- FIELER, A. C., M. ESLAVA, AND D. Y. XU (2018): “Trade, Quality Upgrading, and Input Linkages: Theory and Evidence from Colombia,” *American Economic Review*, 108, 109–46.
- FLINN, C. J. (2006): “Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates,” *Econometrica*, 74, 1013–1062.
- GERARD, F. AND G. GONZAGA (2021): “Informal Labor and the Efficiency Cost of Social Programs: Evidence from Unemployment Insurance in Brazil,” *American Economic Journal: Economic Policy*, 13, 167–206.
- GERVAIS, A. AND J. B. JENSEN (2019): “The Tradability of Services: Geographic Concentration and Trade Costs,” *Journal of International Economics*, 118, 331 – 350.
- GOLDBERG, P. K. AND N. PAVCNİK (2003): “The response of the informal sector to trade liberalization,” *Journal of Development Economics*, 72, 463–496.
- GONZAGA, G., W. F. MALONEY, AND A. MIZALA (2003): “Labor Turnover and Labor Legislation in Brazil,” *Economia*, 4, 165–222.
- HEAD, K. AND T. MAYER (2014): “Gravity Equations: Workhorse, Toolkit, and Cookbook,” Elsevier, vol. 4 of *Handbook of International Economics*, 131–195.
- HECKMAN, J. J. AND C. PAGÉS (2000): “The Cost of Job Security Regulation: Evidence from Latin American Labor Markets,” Working Paper 7773, National Bureau of Economic Research.
- HELPMAN, E., O. ITSKHOKI, M.-A. MUENDLER, AND S. J. REDDING (2017): “Trade and Inequality: From



- Theory to Estimation,” *The Review of Economic Studies*, 84, 357–405.
- HOPENHAYN, H. AND R. ROGERSON (1993): “Job Turnover and Policy Evaluation: A General Equilibrium Analysis,” *Journal of Political Economy*, 101, 915–938.
- HORNSTEIN, A., P. KRUSELL, AND G. L. VIOLANTE (2011): “Frictional Wage Dispersion in Search Models: A Quantitative Assessment,” *American Economic Review*, 101, 2873–98.
- HSIEH, C.-T. AND P. J. KLENOW (2009): “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124, 1403 – 1448.
- LA PORTA, R. AND A. SHLEIFER (2008): “The Unofficial Economy and Economic Development,” *Brooking Papers on Economic Activity*, 105, 473–522.
- (2014): “Informality and Development,” *Journal of Economic Perspectives*, 28, 109–26.
- LJUNGQVIST, L. AND T. J. SARGENT (2017): “The Fundamental Surplus,” *American Economic Review*, 107, 2630–65.
- MCCAIG, B. AND N. PAVCNİK (2018): “Export Markets and Labor Allocation in a Low-Income Country,” *American Economic Review*, 108, 1899–1941.
- MEGHIR, C., R. NARITA, AND J.-M. ROBIN (2015): “Wages and Informality in Developing Countries,” *The American Economic Review*, 105, 1509–1546.
- MELITZ, M. J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71, 1695–1725.
- MORTENSEN, D. T. AND C. A. PISSARIDES (1999): “Unemployment Responses to ‘Skill-Biased’ Technology Shocks: The Role of Labour Market Policy,” *The Economic Journal*, 109, 242–265.
- OLLEY, G. S. AND A. PAKES (1996): “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 64, 1263–1297.
- PAVCNIK, N. (2002): “Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants,” *The Review of Economic Studies*, 69, 245–276.
- PERRY, G., W. MALONEY, O. ARIAS, P. FAJNZYLBER, A. MASON, AND J. SAAVEDRA-CHANDUVI (2007): *Informality: Exit or exclusion*, The World Bank. Washington, D.C.
- PETRONGOLO, B. AND C. A. PISSARIDES (2001): “Looking into the Black Box: A Survey of the Matching Function,” *Journal of Economic Literature*, 39, 390–431.
- PONCZEK, V. AND G. ULYSSEA (2022): “Enforcement of Labour Regulation and the Labour Market Effects of Trade: Evidence from Brazil,” *The Economic Journal*, 132, 361–390.
- SONG, J., D. J. PRICE, F. GUVENEN, N. BLOOM, AND T. VON WACHTER (2019): “Firming up inequality,” *The Quarterly Journal of Economics*, 134, 1–50.
- ULYSSEA, G. (2010): “Regulation of Entry, Labor Market Institutions and the Informal Sector,” *Journal of Development Economics*, 91, 87–99.
- (2018): “Firms, Informality, and Development: Theory and Evidence from Brazil,” *American Economic Review*, 108, 2015–47.
- (2020): “Informality: Causes and Consequences for Development,” *Annual Review of Economics*, 12, 525–546.
- WORLD BANK (2019): *World Development Report 2019 : The Changing Nature of Work*, washington, DC.
- (2022): “World Development Indicators: Structure of demand,” <https://wdi.worldbank.org/table/4.8>, accessed: January 2024.



# **Online Appendix**

## **Trade and Domestic Distortions: The Case of Informality**

by Rafael Dix-Carneiro, Pinelopi Koujianou Goldberg, Costas Meghir, Gabriel Ulyssea

# Table of Contents

<b>A Steady-State Distribution of States</b>	<b>A4</b>
A.1 Informal Firms . . . . .	A4
A.2 Formal Firms . . . . .	A4
<b>B Entry</b>	<b>A6</b>
<b>C Flow Conditions for Workers and Firms</b>	<b>A6</b>
<b>D Vacancies</b>	<b>A9</b>
<b>E Unemployment Benefits / Tax Collection / Transfers</b>	<b>A9</b>
<b>F Service Sector Market Clearing</b>	<b>A10</b>
<b>G Aggregate Income</b>	<b>A10</b>
<b>H Trade Balance</b>	<b>A11</b>
<b>I Worker Value Functions</b>	<b>A11</b>
<b>J Mechanical Effects: Trade Cost or Technology Changes</b>	<b>A13</b>
<b>K Iceberg Trade Cost</b>	<b>A15</b>
<b>L Data Appendix</b>	<b>A15</b>
L.1 RAIS and SECEX . . . . .	A15
L.2 PIA, PAS and PAC (Firm-Level Surveys) and SECEX . . . . .	A17
L.3 ECINF ( <i>Pesquisa de Economia Informal Urbana</i> ) . . . . .	A18
L.4 PME . . . . .	A19
L.5 IBGE National Accounts . . . . .	A20
<b>M Model Fit: Moments Generated by the Model vs. Data</b>	<b>A20</b>
<b>N Background: The Cost of Labor Regulations in Brazil</b>	<b>A23</b>
<b>O Additional Post-Estimation Results: Overlapping Distributions of Productivity</b>	<b>A23</b>
<b>P Additional Post-Estimation Results: Revenue per Worker is Increasing with Firm-Level Total Factor Productivity</b>	<b>A24</b>
<b>Q Sensitivity Analysis</b>	<b>A26</b>
<b>I Estimation</b>	<b>S1</b>
I.1 Estimation Algorithm . . . . .	S1
I.2 Estimation Algorithm – Further Details . . . . .	S7

<b>II Simulation</b>	<b>S18</b>
II.1 Simulation Algorithm . . . . .	S18
II.2 Simulation Algorithm – Details . . . . .	S24

## A Steady-State Distribution of States

### A.1 Informal Firms

Denote by  $G_k(z'|z)$  the cumulative distribution function of  $z'$  conditional on  $z$  and  $g_k(z'|z)$  its density. The period starts with  $N_{ki}$  informal firms and distribution of states  $\psi_{ki}(z, \ell)$  at the very beginning of stage 1. After (endogenous and exogenous) exit, change in formal status, and entry, but before labor adjustment (end of stage 1 / beginning of stage 2) the distribution of states is:

$$\tilde{\psi}_{ki}(z, \ell) \equiv \frac{\mathcal{I}[\ell = 1] M_{ki} \psi_{ki}^e(z) + \mathcal{I}[\ell \geq 1] (1 - \alpha_k) N_{ki} \psi_{ki}(z, \ell) I_{ki}^{stay}(z, \ell)}{N_{ki}} \quad (\text{A.1})$$

$$= \mathcal{I}[\ell = 1] \frac{M_{ki}}{N_{ki}} \psi_{ki}^e(z) + \mathcal{I}[\ell \geq 1] (1 - \alpha_k) \psi_{ki}(z, \ell) I_{ki}^{stay}(z, \ell), \quad (\text{A.2})$$

where  $M_{ki}$  is the mass of informal entrants into sector  $k$ ,  $\psi_{ki}^e(z)$  is the distribution of  $z$  productivities of entrants conditional on entry into the informal sector:

$$\psi_{ki}^e(z) \equiv \frac{g_k^e(z) I_k^{informal}(z)}{\int_{\tilde{z}} g_k^e(\tilde{z}) I_k^{informal}(\tilde{z}) d\tilde{z}}. \quad (\text{A.3})$$

In equation (A.1) the numerator is the total mass of firms with state  $(z, \ell)$ . The denominator is the total mass of firms at the stage we consider. In steady state, entrants replace firms who exit, so that there are  $N_{ki}$  firms at that stage.

After firms make adjustment decisions, and at the production stage (end of stage 2 / beginning of stage 3), the distribution of states is:

$$\hat{\psi}_{ki}(z, \ell') \equiv \int_{\ell} \tilde{\psi}_{ki}(z, \ell) \mathcal{I}[L_{ki}(z, \ell) = \ell'] d\ell. \quad (\text{A.4})$$

At the end of the period, after production takes place, firms draw their productivity  $z'$  for the next period (stage 3). In steady state, the distribution of states at the very end of the period (end of stage 3) replicates the initial one (very beginning of stage 1):

$$\psi_{ki}(z', \ell') = \int_z \hat{\psi}_{ki}(z, \ell') g_k(z'|z) dz. \quad (\text{A.5})$$

To fix ideas, Table A.1 clarifies the notation for the distribution of states in different stages within a period.

### A.2 Formal Firms

The period starts with  $N_{kf}$  formal firms and distribution of states  $\psi_{kf}(z, \ell)$  at the very beginning of stage 1. After (endogenous and exogenous) exit, change in formal status, and entry, but before labor adjustment

Table A.1: Distributions of States at Different Stages

$\psi_{kj}$	Distribution of states at the very beginning, and at the very end of the period—very beginning of stage 1 and very end of stage 3
$\tilde{\psi}_{kj}$	Distribution of states right after entry, exit, and change of formal status but <b>before</b> labor adjustment—very end of stage 1 / very beginning of stage 2
$\hat{\psi}_{kj}$	Distribution of states after labor adjustment, at the production stage—end of stage 2

(end of stage 1 / beginning of stage 2) the distribution of states is:

$$\tilde{\psi}_{kf}(z, \ell) \equiv \frac{\mathcal{I}[\ell = 1] M_{kf} \psi_{kf}^e(z) + \mathcal{I}[\ell \geq 1] (1 - \alpha_k) N_{ki} \psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell) + \mathcal{I}[\ell \geq 1] (1 - \alpha_k) N_{kf} \psi_{kf}(z, \ell) I_{kf}^{stay}(z, \ell)}{N_{kf}} \quad (\text{A.6})$$

$$= \mathcal{I}[\ell = 1] \frac{M_{kf}}{N_{kf}} \psi_{kf}^e(z) + \mathcal{I}[\ell \geq 1] (1 - \alpha_k) \frac{N_{ki}}{N_{kf}} \psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell) + \mathcal{I}[\ell \geq 1] (1 - \alpha_k) \psi_{kf}(z, \ell) I_{kf}^{stay}(z, \ell), \quad (\text{A.7})$$

where  $M_{kf}$  is the mass of formal entrants into sector  $k$ ,  $\psi_{kf}^e(z)$  is the distribution of  $z$  productivities of entrants conditional on entry into the formal sector:

$$\psi_{kf}^e(z) \equiv \frac{g_k^e(z) I_k^{formal}(z)}{\int_{\tilde{z}} g_k^e(\tilde{z}) I_k^{formal}(\tilde{z}) d\tilde{z}}. \quad (\text{A.8})$$

The numerator in equation (A.6) is the total mass of firms with state  $(z, \ell)$ . The denominator is the total mass of firms at the stage we consider. In steady state, entrants replace firms who exit, so that there are  $N_{kf}$  firms at that stage. After firms make adjustment decisions, and at the production stage (end of stage 2 / beginning of stage 3), the distribution of states is:

$$\hat{\psi}_{kf}(z, \ell') \equiv \int_{\ell} \tilde{\psi}_{kf}(z, \ell) \mathcal{I}(I_{kf}(z, \ell) = \ell') d\ell. \quad (\text{A.9})$$

At the end of the period, after production takes place, firms draw their productivity  $z'$  for the next period (stage 3). In steady state, the distribution of states at the very end of the period (very end of stage 3) replicates the initial one (very beginning of stage 1):

$$\psi_{kf}(z', \ell') = \int_z \hat{\psi}_{kf}(z, \ell') g_k(z'|z) dz.$$

## B Entry

Let  $M_k$  denote the mass of entrants in sector  $k = C, S$ . The fraction of entrants into the formal and informal sectors are given respectively by  $\omega_{kf}$  and  $\omega_{ki}$ :

$$\omega_{kf} \equiv \Pr \left( I_k^{formal}(z) = 1 \right) = \int_z g_k^e(z) dz, \quad (\text{B.1})$$

$$\omega_{ki} \equiv \Pr \left( I_k^{informal}(z) = 1 \right) = \int_z g_k^e(z) dz. \quad (\text{B.2})$$

Therefore, the masses of entrants in the formal and informal sectors are given by:

$$M_{ki} = \omega_{ki} M_k, \quad (\text{B.3})$$

$$M_{kf} = \omega_{kf} M_k. \quad (\text{B.4})$$

The masses of entrants into each sector,  $M_k$ , are pinned down by the free entry condition (assuming positive entry in both sectors):

$$c_{e,k} = V_k^e = \int_z \left[ V_{ki}^e(z) I_k^{informal}(z) + V_{kf}^e(z) I_k^{formal}(z) \right] g_k^e(z) dz. \quad (\text{B.5})$$

## C Flow Conditions for Workers and Firms

In order to write the labor market clearing conditions, we first define the following quantities.

- Number of workers at the **beginning** of the period in sector  $k$  (before entry, exit, change of formal status and labor adjustment), working in formal or informal firms ( $T$  stands for "total"):

$$W_{kj}^T = N_{kj} \underbrace{\int_z \int_\ell \ell \psi_{kj}(z, \ell) d\ell dz}_{\text{avg. \# of workers per firm}} = L_{kj} \quad (\text{C.1})$$

for  $j = f, i$  and  $k = C, S$ .

- Number of workers in sector  $(k, j)$  who are fired because their firms receive a destruction shock:

$$W_{kj}^{DS} = \alpha_{kj} N_{kj} \int_z \int_\ell \ell \psi_{kj}(z, \ell) d\ell dz = \alpha_{kj} L_{kj} \quad (\text{C.2})$$

- Number of workers in sector  $(k, j)$  who are fired due to **endogenous** firm exit:

$$W_{kj}^{EE} = (1 - \alpha_{kj}) N_{kj} \times \int_z \int_\ell \ell \psi_{kj}(z, \ell) I_{kj}^{exit}(z, \ell) d\ell dz \quad (\text{C.3})$$

where  $(1 - \alpha_{kj}) N_{kj}$  is the mass of firms that survive after the destruction shock hits.

- Number (mass) of **surviving incumbent** firms in sector  $(k, j)$  in the interim period:

$$N'_{kj} \equiv (1 - \alpha_{kj}) N_{kj} \int_z \int_\ell \psi_{kj}(z, \ell) I_{kj}^{stay}(z, \ell) d\ell dz \quad (\text{C.4})$$

- Number of workers initially in sector  $(k, j)$  who are fired due to downsizing at the interim stage:

$$W_{kj}^D = N'_{kj} \int_z \int_\ell \tilde{\psi}_{kj}^{incumbent}(z, \ell) (1 - I_{kj}^{hire}(z, \ell)) (\ell - L_{kj}(z, \ell)) d\ell dz \quad (C.5)$$

where  $\tilde{\psi}_{kj}^{incumbent}(z, \ell)$  is the distribution of states in the interim stage among **surviving incumbents**. Note that this is not the same distribution as  $\tilde{\psi}_{kj}(z, \ell)$  as it does not include entrants. It is obtained as follows:

$$\begin{aligned} \tilde{\psi}_{kj}^{incumbent}(z, \ell) &\equiv \frac{(1 - \alpha_{kj}) N_{kj} \psi_{kj}(z, \ell) I_{kj}^{stay}(z, \ell)}{N'_{kj}} \\ &= \frac{\psi_{kj}(z, \ell) I_{kj}^{stay}(z, \ell)}{\int_{\tilde{z}} \int_{\tilde{\ell}} \psi_{kj}(\tilde{z}, \tilde{\ell}) I_{kj}^{stay}(\tilde{z}, \tilde{\ell}) d\tilde{\ell} d\tilde{z}} \end{aligned} \quad (C.6)$$

- Total fraction of workers in the formal sector of sector  $k$  who are laid off, conditional on starting the period in a formal firm in sector  $k$ :

$$\begin{aligned} \chi_{kf}^{layoff} &= \frac{W_{kf}^{DS} + W_{kf}^{EE} + W_{kf}^D}{W_{kf}^T} \\ &= \alpha_k + \frac{\left( \begin{aligned} &(1 - \alpha_k) \int_z \int_\ell \ell \psi_{kf}(z, \ell) I_{kf}^{exit}(z, \ell) d\ell dz \\ &+ (1 - \alpha_k) \left( \int_z \int_\ell \psi_{kf}(z, \ell) I_{kf}^{stay}(z, \ell) d\ell dz \right) \times \\ &\int_z \int_\ell \tilde{\psi}_{kf}^{incumbent}(z, \ell) (1 - I_{kf}^{hire}(z, \ell)) (\ell - L_{kf}(z, \ell)) d\ell dz \end{aligned} \right)}{\int_z \int_\ell \ell \psi_{kf}(z, \ell) d\ell dz} \end{aligned} \quad (C.7)$$

- Number of firms that start the period as informal firms, but end the period as formal firms (because they formalized).

$$N'_{ki \rightarrow f} \equiv (1 - \alpha_k) N_{ki} \int_z \int_\ell \psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell) d\ell dz, \quad (C.8)$$

where  $(1 - \alpha_k) N_{ki}$  is the mass of firms that survive after the destruction shock hits.

- Distribution of states among firms that switched from informal to formal, in the interim period—before adjusting the labor force.

$$\begin{aligned} \tilde{\psi}_{ki \rightarrow f}(z, \ell) &\equiv \frac{(1 - \alpha_k) N_{ki} \psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell)}{N'_{ki \rightarrow f}} \\ &= \frac{\psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell)}{\int_{\tilde{z}} \int_{\tilde{\ell}} \psi_{ki}(\tilde{z}, \tilde{\ell}) I_{ki}^{change}(\tilde{z}, \tilde{\ell}) d\tilde{\ell} d\tilde{z}}. \end{aligned} \quad (C.9)$$

- Number of workers who start the period in informal firms, but end the period in formal firms (their employers switched to formal, and they were not fired after the interim productivity was realized):

$$W_{k, i \rightarrow f} = N'_{ki \rightarrow f} \int_z \int_\ell \tilde{\psi}_{ki \rightarrow f}(z, \ell) \left( \begin{aligned} &\ell \times I_{kf}^{hire}(z, \ell) + \\ &L_{kf}(z, \ell) \times (1 - I_{kf}^{hire}(z, \ell)) \end{aligned} \right) d\ell dz \quad (C.10)$$

- Fraction of workers who start the period in informal firms, but end the period in formal firms:

$$\begin{aligned}\chi_{ki \rightarrow f}^{change} &= \frac{W_{k,i \rightarrow f}}{W_{ki}^T} \\ &= \frac{\left( (1 - \alpha_k) \left( \int_z \int_\ell \psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell) d\ell dz \right) \times \right.}{\int_z \int_\ell \ell \psi_{ki}(z, \ell) d\ell dz} \\ &\quad \left. \int_z \int_\ell \tilde{\psi}_{ki \rightarrow f}(z, \ell) \left( \ell \times I_{kf}^{hire}(z, \ell) + L_{kf}(z, \ell) \times (1 - I_{kf}^{hire}(z, \ell)) \right) d\ell dz \right) \end{aligned} \quad (C.11)$$

- Number of workers who start the period in informal firms, but their employers switched to formal status:

$$W_{ki}^{SF} = (1 - \alpha_k) N_{ki} \int_z \int_\ell \ell \psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell) d\ell dz \quad (C.12)$$

- Fraction of workers who start employed in the informal sector and leave it in the interim period (became unemployed or employer switched to formal):

$$\begin{aligned}\chi_{ki}^{leave} &= \frac{W_{ki}^{DS} + W_{ki}^{EE} + W_{ki}^{SF} + W_{ki}^D}{W_{ki}^T} \\ &= \alpha_k + \frac{\left( (1 - \alpha_k) \int_z \int_\ell \ell \psi_{ki}(z, \ell) I_{ki}^{exit}(z, \ell) d\ell dz + \right.}{\int_z \int_\ell \ell \psi_{ki}(z, \ell) d\ell dz} \\ &\quad \left( (1 - \alpha_k) \int_z \int_\ell \ell \psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell) d\ell dz + \right. \\ &\quad \left. (1 - \alpha_k) \left( \int_z \int_\ell \psi_{ki}(z, \ell) I_{ki}^{stay}(z, \ell) d\ell dz \right) \times \right. \\ &\quad \left. \left. \int_z \int_\ell \tilde{\psi}_{ki}^{incumbent}(z, \ell) (1 - I_{ki}^{hire}(z, \ell)) (\ell - L_{ki}(z, \ell)) d\ell dz \right) \right) \end{aligned} \quad (C.13)$$

With these objects, we can define the equilibrium conditions that refer to labor market flows:

$$\chi_{ki}^{leave} L_{ki} = L_u \mu_{ki}^e \quad (C.14)$$

$$\chi_{kf}^{layoff} L_{kf} = L_u \mu_{kf}^e + L_{ki} \chi_{ki \rightarrow f}^{change}. \quad (C.15)$$

These conditions state that the mass of workers in each sector  $(k, j)$  cannot be contracting or expanding in equilibrium (expressions (C.14) and (C.15)). Finally, the sum of unemployment and employment levels across sectors equals the total labor force  $\bar{L}$ :

$$L_{Cf} + L_{Ci} + L_{Sf} + L_{Si} + L_u = \bar{L}. \quad (C.16)$$

We can proceed in a similar way to define the equilibrium flow conditions for firms. The relevant objects follow.

- Fraction of formal firms exiting sector  $k$ :

$$\varrho_{kf}^{exit} = \alpha_k + (1 - \alpha_k) \int_z \int_\ell I_{kf}^{exit}(z, \ell) \psi_{kf}(z, \ell) d\ell dz \quad (C.17)$$

- Fraction of informal firms exiting sector  $k$ :

$$\varrho_{ki}^{exit} = \alpha_k + (1 - \alpha_k) \int_z \int_\ell \left( I_{ki}^{exit}(z, \ell) + I_{ki}^{change}(z, \ell) \right) \psi_{ki}(z, \ell) d\ell dz \quad (C.18)$$



- Fraction of informal firms changing status in sector  $k$ :

$$\varrho_{ki}^{change} = (1 - \alpha_k) \int_z \int_\ell I_{ki}^{change}(z, \ell) \psi_{ki}(z, \ell) d\ell dz \quad (C.19)$$

Similarly to workers, the mass of firms in each sector  $(k, j)$  must be constant in steady state. This means that the inflow of firms must equal the outflow, which can be written as:

$$\varrho_{kf}^{exit} N_{kf} = M_{kf} + \varrho_{ki}^{change} N_{ki}, \quad (C.20)$$

$$\varrho_{ki}^{exit} N_{ki} = M_{ki}. \quad (C.21)$$

## D Vacancies

Aggregate vacancies in sector  $kj$  are given by:

$$V_{kj} = N_{kj} \int_z \int_\ell v_{kj}(z, \ell) \tilde{\psi}_{ki}(z, \ell) d\ell dz + \frac{M_{kj}}{\mu^v} \quad (D.1)$$

where  $v_{kj}(z, \ell)$  is the number of vacancies a firm with productivity  $z$  and labor force  $\ell$  posts and  $\frac{M_{kj}}{\mu^v}$  is the number of vacancies posted at entry (and before adjustment in stage 2).

## E Unemployment Benefits / Tax Collection / Transfers

Government Revenue is given by the sum of value-added taxes, payroll taxes, firing costs and import taxes:

$$\begin{aligned} G_{Rev} = & \sum_k N_{kf} \tau_y \int_z \int_{\ell'} V A_k(z, \ell') \hat{\psi}_{kf}(z, \ell') d\ell' dz \\ & + \sum_k N_{kf} \tau_w \int_z \int_{\ell'} \max\{w_{kf}(z, \ell'), \underline{w}\} \ell \hat{\psi}_{kf}(z, \ell') d\ell' dz \\ & + \sum_k N_{kf} \kappa \int_z \int_\ell \tilde{\psi}_{kf}(z, \ell) (\ell - L_{kf}(z, \ell)) (1 - I_{kf}^{hire}(z, \ell)) d\ell dz \\ & + (\tau_a - 1) \frac{D_{H,C} (\epsilon \tau_a \tau_c)^{1-\sigma_C}}{\tau_a}. \end{aligned} \quad (E.1)$$

Government spending with unemployment insurance is given by:

$$G_{UI} = b^u \times \underbrace{\sum_k (W_{kf}^{DS} + W_{kf}^{EE} + W_{kf}^D)}_{\text{mass of formal workers who transition to unemployment}} \quad (E.2)$$

We impose that

$$T = G_{Rev} - G_{UI} \geq 0 \quad (E.3)$$

and that  $T$  is rebated to consumers.

Important note: part of the aggregate informality costs

$$\sum_k N_{ki} \int_z \int_{\ell'} p_{ki}(\ell') R_k(z, \ell') \hat{\psi}_{ki}(z, \ell') d\ell' dz \quad (E.4)$$

should be considered government revenue as these consist of fines. However, part of these costs should not, as they consist of opportunity costs associated with informality. Therefore, we do not add these costs to government revenue. However, the model redistributes these costs to consumers. One way to view this procedure is that these costs affect/distort the decisions of firms, but we do not consider these costs as wasted resources.

## F Service Sector Market Clearing

Service sector goods are used for final consumption (consumers spend  $(1 - \zeta) I$  on it), intermediate inputs (firms spend  $X_S^{int}$  on it) and as inputs for hiring costs, fixed costs and entry costs (and fixed costs of exporting). The average (per firm) hiring costs in sector  $(k, j)$ :

$$\bar{H}_{kj} = \int_z \int_\ell H_{kj}(\ell, L_{kj}(z, \ell)) I_{kj}^{hire}(z, \ell) \tilde{\psi}_{kj}(z, \ell) d\ell dz, \quad (\text{F.1})$$

and the fraction of manufacturing-sector goods firms that export is given by:

$$\mu_x = \int_z \int_{\ell'} \hat{\psi}_{Cf}(z, \ell') I_C^x(z, \ell') d\ell' dz \quad (\text{F.2})$$

Expenditure on entry and hiring costs, fixed costs of operations and export costs are given by:

$$E_S = \sum_{k=C, S; j=i, f} N_{kj} (\bar{H}_{kj} + \bar{c}_{kj}) + N_{Cf} \mu_x f_x + \sum_{k=C, S} M_k c_{e, k} \quad (\text{F.3})$$

## G Aggregate Income

Aggregate income is given by total wages, government transfers and total profits:

$$\begin{aligned} I = & \sum_k N_{ki} \int_z \int_{\ell'} w_{ki}(z, \ell') \ell' \hat{\psi}_{ki}(z, \ell') d\ell' dz \\ & + \sum_k N_{kf} \int_z \int_{\ell'} \max\{w_{kf}(z, \ell'), \underline{w}\} \ell' \hat{\psi}_{kf}(z, \ell') d\ell' dz \\ & + \sum_k N_{ki} \int_z \int_{\ell'} \tilde{\pi}_{ki}(z, \ell') \hat{\psi}_{ki}(z, \ell') d\ell' dz \\ & + \sum_k N_{kf} \int_z \int_{\ell'} \tilde{\pi}_{kf}(z, \ell') \hat{\psi}_{kf}(z, \ell') d\ell' dz \\ & + G_{Rev} \\ & + \sum_k N_{ki} \int_z \int_{\ell'} p_{ki}(\ell') R_k(z, \ell') \hat{\psi}_{ki}(z, \ell') d\ell' dz \\ & - \sum_k N_{kf} \kappa \int_z \int_\ell \tilde{\psi}_{kf}(z, \ell) (\ell - L_{kf}(z, \ell)) (1 - I_{kf}^{hire}(z, \ell)) d\ell dz \\ & - \sum_{k=C, S; j=i, f} N_{kj} \bar{H}_{kj} \\ & - \sum_k M_k c_{e, k}, \end{aligned} \quad (\text{G.1})$$

where profits  $\tilde{\pi}$  are computed **before** subtracting hiring costs.

## H Trade Balance

Trade balance implies that total imports must equal total exports, which is given by:

$$\frac{D_{H,C} (\epsilon \tau_a \tau_c)^{1-\sigma_C}}{\tau_a} = Exports \quad (\text{H.1})$$

## I Worker Value Functions

**Present value of a formal job at a firm with state  $(z, \ell')$  at the production stage**

$$\begin{aligned} \bar{J}_{kf}^e(z, \ell') &= w_{kf}(z, \ell') \\ &+ \frac{1 - \alpha_k}{1 + r} E_{z'|z} \left( \begin{aligned} &\left( \frac{\alpha_k}{1 - \alpha_k} + I_{kf}^{exit}(z', \ell') \right) \times \left( b + b_u + \frac{1}{1+r} J^u \right) \\ &+ I_{kf}^{stay}(z', \ell') \times I_{kf}^{contract}(z', \ell') \times p_{kf}^{fire}(z', \ell') \times \left( b + b_u + \frac{1}{1+r} J^u \right) \\ &+ I_{kf}^{stay}(z', \ell') \times I_{kf}^{contract}(z', \ell') \times \left( 1 - p_{kf}^{fire}(z', \ell') \right) \times \bar{J}_{kf}^e(z', L_{kf}(z', \ell')) \\ &+ I_{kf}^{stay}(z', \ell') \times I_{kf}^{expand}(z', \ell') \times \bar{J}_{kf}^e(z', L_{kf}(z', \ell')) \end{aligned} \right) \\ p_{kf}^{fire}(z', \ell') &\equiv \frac{\ell' - L_{kf}(z', \ell')}{\ell'} \\ I_{kf}^{contract}(z', \ell') &\equiv I(L_{kf}(z', \ell') < \ell') \\ I_{kf}^{expand}(z', \ell') &\equiv I(L_{kf}(z', \ell') \geq \ell') \end{aligned}$$

Rewriting:

$$\begin{aligned} \bar{J}_{kf}^e(z, \ell') &= w_{kf}(z, \ell') \\ &+ \frac{1 - \alpha_k}{1 + r} \left( \frac{\alpha_k}{1 - \alpha_k} + E_{z'|z} [I_{kf}^{exit}(z', \ell')] \right) \times \left( b + b_u + \frac{1}{1+r} J^u \right) \\ &+ \frac{1 - \alpha_k}{1 + r} E_{z'|z} \left[ I_{kf}^{stay}(z', \ell') \times I_{kf}^{contract}(z', \ell') \times p_{kf}^{fire}(z', \ell') \right] \times \left( b + b_u + \frac{1}{1+r} J^u \right) \\ &+ \frac{1 - \alpha_k}{1 + r} E_{z'|z} \left[ I_{kf}^{stay}(z', \ell') \times I_{kf}^{contract}(z', \ell') \times \left( 1 - p_{kf}^{fire}(z', \ell') \right) \times \bar{J}_{kf}^e(z', L_{kf}(z', \ell')) \right] \\ &+ \frac{1 - \alpha_k}{1 + r} E_{z'|z} \left[ I_{kf}^{stay}(z', \ell') \times I_{kf}^{expand}(z', \ell') \times \bar{J}_{kf}^e(z', L_{kf}(z', \ell')) \right] \end{aligned}$$

It will be convenient to work with:

$$J_{kf}^e(z, \ell') \equiv (1 + r) \left( \bar{J}_{kf}^e(z, \ell') - w_{kf}(z, \ell') \right) \quad (\text{I.1})$$

**Present value of an informal job at a firm with state  $(z, \ell')$  at the production stage**

$$\begin{aligned}
\bar{J}_{ki}^e(z, \ell') &= w_{ki}(z, \ell') \\
&+ \frac{(1 - \alpha_k)}{1 + r} E_{z'|z} \left( \begin{aligned} &\left( \frac{\alpha_k}{(1 - \alpha_k)} + I_{ki}^{exit}(z', \ell') \right) \times \left( b + \frac{1}{1+r} J^u \right) \\ &+ I_{ki}^{stay}(z', \ell') \times I_{ki}^{contract}(z', \ell') \times p_{ki}^{fire}(z', \ell') \times \left( b + \frac{1}{1+r} J^u \right) \\ &+ I_{ki}^{stay}(z', \ell') \times I_{ki}^{contract}(z', \ell') \times \left( 1 - p_{ki}^{fire}(z', \ell') \right) \times \bar{J}_{ki}^e(z', L_{ki}(z', \ell')) \\ &+ I_{ki}^{stay}(z', \ell') \times I_{ki}^{expand}(z', \ell') \times \bar{J}_{ki}^e(z', L_{ki}(z', \ell')) \\ &+ I_{ki}^{change}(z', \ell') \times I_{kf}^{contract}(z', \ell') \times p_{kf}^{fire}(z', \ell') \times \left( b + b_u + \frac{1}{1+r} J^u \right) \\ &+ I_{ki}^{change}(z', \ell') \times I_{kf}^{contract}(z', \ell') \times \left( 1 - p_{kf}^{fire}(z', \ell') \right) \times \bar{J}_{kf}^e(z', L_{kf}(z', \ell')) \\ &+ I_{ki}^{change}(z', \ell') \times I_{kf}^{expand}(z', \ell') \times \bar{J}_{kf}^e(z', L_{kf}(z', \ell')) \end{aligned} \right) \\
p_{ki}^{fire}(z', \ell') &\equiv \frac{\ell' - L_{ki}(z', \ell')}{\ell'} \\
I_{ki}^{contract}(z', \ell') &\equiv I(L_{ki}(z', \ell') < \ell') \\
I_{ki}^{expand}(z', \ell') &\equiv I(L_{ki}(z', \ell') \geq \ell')
\end{aligned}$$

Rewriting

$$\begin{aligned}
\bar{J}_{ki}^e(z, \ell') &= w_{ki}(z, \ell') \\
&+ \frac{(1 - \alpha_k)}{1 + r} \left( \frac{\alpha_k}{(1 - \alpha_k)} + E_{z'|z} [I_{ki}^{exit}(z', \ell')] \right) \times \left( b + \frac{1}{1+r} J^u \right) \\
&+ \frac{(1 - \alpha_k)}{1 + r} E_{z'|z} \left[ I_{ki}^{stay}(z', \ell') \times I_{ki}^{contract}(z', \ell') \times p_{ki}^{fire}(z', \ell') \right] \times \left( b + \frac{1}{1+r} J^u \right) \\
&+ \frac{(1 - \alpha_k)}{1 + r} E_{z'|z} \left[ I_{ki}^{stay}(z', \ell') \times I_{ki}^{contract}(z', \ell') \times \left( 1 - p_{ki}^{fire}(z', \ell') \right) \times \bar{J}_{ki}^e(z', L_{ki}(z', \ell')) \right] \\
&+ \frac{(1 - \alpha_k)}{1 + r} E_{z'|z} \left[ I_{ki}^{stay}(z', \ell') \times I_{ki}^{expand}(z', \ell') \times \bar{J}_{ki}^e(z', L_{ki}(z', \ell')) \right] \\
&+ \frac{(1 - \alpha_k)}{1 + r} E_{z'|z} \left[ I_{ki}^{change}(z', \ell') \times I_{kf}^{contract}(z', \ell') \times p_{kf}^{fire}(z', \ell') \right] \times \left( b + b_u + \frac{1}{1+r} J^u \right) \\
&+ \frac{(1 - \alpha_k)}{1 + r} E_{z'|z} \left[ I_{ki}^{change}(z', \ell') \times I_{kf}^{contract}(z', \ell') \times \left( 1 - p_{kf}^{fire}(z', \ell') \right) \times \bar{J}_{kf}^e(z', L_{kf}(z', \ell')) \right] \\
&+ \frac{(1 - \alpha_k)}{1 + r} E_{z'|z} \left[ I_{ki}^{change}(z', \ell') \times I_{kf}^{expand}(z', \ell') \times \bar{J}_{kf}^e(z', L_{kf}(z', \ell')) \right]
\end{aligned}$$

As before, it will be convenient to work with:

$$J_{ki}^e(z, \ell') \equiv (1 + r) \left( \bar{J}_{ki}^e(z, \ell') - w_{ki}(z, \ell') \right) \quad (I.2)$$

**Value of Search**

$$J^u = \sum_{k,j} \mu_{kj}^e \int_{\ell} \int_{z'} \bar{J}_{kj}^e(z, L_{kj}(z, \ell)) g_{kj}(z, \ell) dz d\ell + \left( 1 - \sum_{k,j} \mu_{kj}^e \right) \left( b + \frac{1}{1+r} J^u \right) \quad (I.3)$$

$$\begin{aligned}\tilde{g}_{kj}(z, \ell) &= N_{kj} \tilde{\psi}_{kj}(z, \ell) v_{kj}(z, \ell) + \mathcal{I}[\ell = 1] \frac{M_{kj}}{\mu^v} \psi_{kj}^e(z) \\ g_{kj}(z, \ell) &= \frac{\tilde{g}_{kj}(z, \ell)}{\int_z \int_\ell \tilde{g}_{kj}(z, \ell) d\ell dz} = \frac{\tilde{g}_{kj}(z, \ell)}{V_{kj}}\end{aligned}$$

## J Mechanical Effects: Trade Cost or Technology Changes

A change in trade costs has two “mechanical” effects: (a) the producer of the sector  $C$  composite good has access to cheaper foreign  $C$  sector differentiated varieties, directly reducing both the prices of the  $C$  sector composite good and the final consumption good; and (b) differentiated variety producers in both sectors have access to cheaper intermediate inputs, reducing their costs and allowing them to charge lower prices. These effects will be materialized even in the absence of general equilibrium and reallocation effects. Our subsequent derivations measure these mechanical effects. In what follows, we fix the allocations, export decisions, wages and aggregate income at the Benchmark level.

The price that a differentiated variety producer in sector  $k \in \{C, S\}$  charges as a function of quantity  $q$  sold is given by:

$$p_k(q) = \left( \frac{X_k}{P_k^{1-\sigma_k}} \right)^{\frac{1}{\sigma_k}} q^{-\frac{1}{\sigma_k}}.$$

Following a change in trade costs, and for fixed demand shifter  $\exp(d_{H,k}) = \left( \frac{X_k}{P_k^{1-\sigma_k}} \right)^{\frac{1}{\sigma_k}}$  for  $k \in \{C, S\}$ , firms will be able to sell larger quantities as they can now charge lower prices (due to lower costs of intermediates). The optimal quantity of the intermediate input bundle in sector  $k \in \{C, S\}$  by a firm of type  $j \in \{i, f\}$  is given by:

$$\iota_{kj}(z, \ell) = \frac{(1 - \delta_k)(\sigma_k - 1) R_{kj}(z, \ell)}{\sigma_k P_k^m},$$

where revenue is given by:

$$R_{kj}(z, \ell) = \exp(d_{H,k} + d_F \mathcal{I}_{kj}^x(z, \ell)) \left( z \ell^{\delta_k} \iota_{kj}(z, \ell)^{1-\delta_k} \right)^{\frac{\sigma_k - 1}{\sigma_k}}.$$

Substituting yields:

$$\iota_{kj}(z, \ell) = \left( \frac{(1 - \delta_k)(\sigma_k - 1) \exp(d_{H,k} + d_F \mathcal{I}_{kj}^x(z, \ell)) (z \ell^{\delta_k})^{\frac{\sigma_k - 1}{\sigma_k}}}{\sigma_k P_k^m} \right)^{\frac{\sigma_k - 1}{\sigma_k - 1} \Lambda_k},$$

where

$$\Lambda_k \equiv \frac{\sigma_k - 1}{\sigma_k - (1 - \delta_k)(\sigma_k - 1)}.$$

We can now write prices:

$$\begin{aligned}p_{kj}(z, \ell) &= \frac{R_{kj}(z, \ell)}{q_{kj}(z, \ell)} \\ &= \exp(d_{H,k} + d_F \mathcal{I}_{kj}^x(z, \ell)) \left( z \ell^{\delta_k} \iota_{kj}(z, \ell)^{1-\delta_k} \right)^{-\frac{1}{\sigma_k}}.\end{aligned}$$

This implies:

$$p_{kj}(z, \ell)^{1-\sigma_k} = (P_k^m)^{-(1-\delta_k)\Lambda_k} \Delta_{kj}(z, \ell),$$

where

$$\Delta_{kj}(z, \ell) = \left( \frac{(1-\delta_k)(\sigma_k-1)}{\sigma_k} \right)^{(1-\delta_k)\Lambda_k} \times \exp \left( ((1-\delta_k)\Lambda_k + (1-\sigma_k)) (d_{H,k} + d_F \mathcal{T}_{kj}^x(z, \ell)) \right) (z\ell^{\delta_k})^{\Lambda_k}.$$

We have that the **domestic** price index in sector  $C$  is given by:

$$P_{C,H}^{1-\sigma_C} = \int \int p_{Cf}(z, \ell)^{1-\sigma_C} N_{Cf} \hat{\psi}_{Cf}(z, \ell) dz d\ell + \int \int p_{Ci}(z, \ell)^{1-\sigma_C} N_{Ci} \hat{\psi}_{Ci}(z, \ell) dz d\ell.$$

And the price index in sector  $S$  is given by:

$$P_S^{1-\sigma_S} = \int \int p_{Sf}(z, \ell)^{1-\sigma_S} N_{Sf} \hat{\psi}_{Sf}(z, \ell) dz d\ell + \int \int p_{Si}(z, \ell)^{1-\sigma_S} N_{Si} \hat{\psi}_{Si}(z, \ell) dz d\ell.$$

And so we end up with the following system:

$$P_C^{1-\sigma_C} = P_{C,H}^{1-\sigma_C} + (\epsilon \tau_a \tau_c)^{1-\sigma_C} \quad (\text{J.1})$$

$$P_{C,H}^{1-\sigma_C} = \tilde{\Omega}_C (P_C^m)^{-(1-\delta_C)\Lambda_C} \quad (\text{J.2})$$

$$P_S^{1-\sigma_S} = \tilde{\Omega}_S (P_S^m)^{-(1-\delta_S)\Lambda_S} \quad (\text{J.3})$$

$$P_k^m = \frac{P_C^{\lambda_k} P_S^{1-\lambda_k}}{\lambda_k^{\lambda_k} (1-\lambda_k)^{1-\lambda_k}} \quad (\text{J.4})$$

$$\tilde{\Omega}_C \equiv \int \int \Delta_{Cf}(z, \ell) N_{Cf} \hat{\psi}_{Cf}(z, \ell) dz d\ell + \int \int \Delta_{Ci}(z, \ell) N_{Ci} \hat{\psi}_{Ci}(z, \ell) dz d\ell \quad (\text{J.5})$$

$$\tilde{\Omega}_S \equiv \int \int \Delta_{Sf}(z, \ell) N_{Sf} \hat{\psi}_{Sf}(z, \ell) dz d\ell + \int \int \Delta_{Si}(z, \ell) N_{Si} \hat{\psi}_{Si}(z, \ell) dz d\ell \quad (\text{J.6})$$

Now, consider a trade cost shock  $\tau_c \rightarrow \tau'_c$ . In that case,  $\tilde{\Omega}_C$  and  $\tilde{\Omega}_S$  are treated as constants in the above system, where  $\Delta_{kj}(z, \ell) = \Delta_{kj}^B(z, \ell)$ , where  $B$  indexes Benchmark equilibrium values.

$$\Delta_{kj}^B(z, \ell) = \left( \frac{(1-\delta_k)(\sigma_k-1)}{\sigma_k} \right)^{(1-\delta_k)\Lambda_k} \times \exp \left( ((1-\delta_k)\Lambda_k + (1-\sigma_k)) \left( d_{H,k}^B + d_F^B \mathcal{T}_{kj}^{x,B}(z, \ell) \right) \right) (z\ell^{\delta_k})^{\Lambda_k}$$

That is, we shock trade costs, but fix allocations, the exchange rate, and demand parameters at their Benchmark levels. As  $\tau_c$  is reduced, this generates a direct effect on  $P_C$  given by equation J.1. This then feeds back into  $P_S$  via  $P_S^m$ , which then feeds back into  $P_C$  via  $P_C^m$ , and so on and so forth.

Similarly, consider a productivity shock  $z \rightarrow \varsigma z$ . Allocations, wages, export decisions and domestic and foreign demand are fixed at Benchmark values. In that case,  $\tilde{\Omega}_C$  and  $\tilde{\Omega}_S$  are functions of  $\varsigma$ :  $\Delta_{kj}(z, \ell) =$

$\Delta_{kj}^\varsigma(z, \ell)$ . Here we shock productivity, but again, we keep allocations and demand parameters fixed.

$$\Delta_{kj}^\varsigma(z, \ell) = \left( \frac{(1 - \delta_k)(\sigma_k - 1)}{\sigma_k} \right)^{(1 - \delta_k)\Lambda_k} \times \exp \left( ((1 - \delta_k)\Lambda_k + (1 - \sigma_k)) \left( d_{H,k}^B + d_F^B \mathcal{I}_{kj}^{x,B}(z, \ell) \right) \right) (\varsigma z \ell^{\delta_k})^{\Lambda_k}$$

In this case,  $\tilde{\Omega}_C$  and  $\tilde{\Omega}_S$  are impacted by the shock. This leads to direct effects on  $P_{C,H}$  and  $P_S$ , which then feed back into each other through linkages.

## K Iceberg Trade Cost

We compute trade costs based on the Head and Ries index below (Head and Mayer, 2014):

$$\phi_{ni} \equiv \sqrt{\frac{X_{ni}X_{in}}{X_{nn}X_{ii}}}.$$

Using the 2003 World Input-Output Database (WIOD),  $i = \textit{Rest of the World}$ , and  $n = \textit{Brazil}$ , we obtain  $\phi_{ni} = 0.02513$ .

Parameterizing the index as:

$$\phi_{ni} = [(1 + t_{ni})\tau_c]^\epsilon$$

Where  $t_{ni}$  is the average tariff between Brazil and the Rest of the World ( $\tau_a - 1 = 12\%$ ). For the trade elasticity, we use  $\epsilon = -3.2$  (median) and  $\epsilon = -4.5$  (mean)—see Table 3.5 of Head and Mayer (2014), who provide a survey of such estimates. This leads to  $\tau_c = 2.82$  and  $\tau_c = 2.02$ , respectively. The average of these two numbers implies  $\tau_c = 2.42$ . We subsequently impose  $\tau_c = 2.4$  in our new analyses.

## L Data Appendix

We use six firm-level datasets containing information on formal and informal firms, as well on their workers. In addition to those, we use one worker-level dataset—*Pesquisa Mensal de Emprego* (PME)—which provides information on workers' allocations and labor market flows. We impose the following common filters across all datasets: we exclude firms and workers in the public sector, agriculture, mining, coal, oil and gas industries. 2003 is our reference year as the ECINF survey is only available for 1997 and 2003. All monetary values (e.g. revenues and wages) correspond to annual values. Finally, we rely on data from the 2000 and 2005 IBGE National Accounts to estimate utility and production function parameters. Sector  $C$  includes all manufacturing sectors (excluding mining, coal, oil and gas industries, as mentioned above). Sector  $S$  includes all services, commerce, construction, transportation, and utilities sectors. In the following sections, we describe the main variables we generate, as well as the moments and auxiliary models computed from each dataset.

### L.1 RAIS and SECEX

RAIS (*Relação Anual de Informações Sociais*) is a matched employer-employee dataset assembled by the Brazilian Ministry of Labor every year since 1976. Establishments are identified by their *Cadastro Nacional*



*de Pessoas Jurídicas* (CNPJ) number, which consists of 14 digits. To make RAIS data compatible with firm-level Census data (PIA, PAS, PAC), we aggregate establishments to the firm level using the first 8 digits of the CNPJ identifier. For multi-establishment firms featuring multiple 4-digit *CNAE* industry codes, we select the code accounting for the largest share of employment within the firm. A negligible share of firms (0.01 percent) have missing industry codes, so they are dropped from the analysis. Firm-level wages and employment are measured as of December of each year. December wages are subsequently annualized. We generate the following firm-level variables:

- *Exit indicator*: We pool RAIS data from 2003 through 2005 to create an exit indicator, which equals one if the firm operates in 2003 but is not found in the data in 2004 nor in 2005.
- *Firm-level employment*: the firm's number of employees, measured in December of each year. Let  $\ell_{i,t}$  denote firm  $i$ 's employment size in year  $t$ .
- *Average firm-level wage*: the firm's annual wage bill divided by number of employees, both measured in December of each year.
- *Firm-level Labor Turnover Rate*: for every firm  $i$ , we define

$$Turnover_i = \frac{|\ell_{i,2004} - \ell_{i,2003}|}{0.5 \times (\ell_{i,2004} + \ell_{i,2003})}.$$

SECEX (*Secretaria de Comércio Exterior*) is an administrative dataset from the federal government containing information on all export and import transactions. These transactions are identified at the firm-level (through the 8 first digits of the CNPJ identifier) and can be merged to the firm-level RAIS data. This procedure allows us to compute exporter indicators for all  $C$ -sector firms. This dummy variable equals one if the firm reports any export transaction in 2003 and zero otherwise (i.e. the firm is found in RAIS but not in SECEX). Using RAIS and SECEX, we compute the following moments and auxiliary models.

***Exit Rate (Formal Firms)*** – see Table M.2

Separately for  $C$ - and  $S$ -sector firms, we compute the mean of the exit dummy variable across all firms.

***Exit Regressions (Formal Firms)*** – see Table M.2

We estimate the following regressions separately for  $C$ - and  $S$ -sector firms:

$$Exit_i = \alpha_k + \beta_k \log(\ell_i) + u_i$$

where  $i$  denotes a firm,  $k = C, S$  denotes sector,  $u_i$  is the error term, and  $Exit_i$  indicates whether firm  $i$ , active in 2003, exits the market in 2004.

***Average Turnover (Formal Firms)*** – see Table M.2

We compute, separately for  $C$ - and  $S$ -sector firms, mean turnover rates across all firms.

***Turnover Regressions*** – see Table M.2

We separately estimate the following regressions, conditional on  $C$ - and  $S$ -sector firms, respectively:

$$\begin{aligned} Turnover_i &= \alpha_C + \beta_C \log(\ell_{i,2003}) + \gamma_C Exporter_{i,2003} + u_i \\ Turnover_i &= \alpha_S + \beta_S \log(\ell_{i,2003}) + u_i \end{aligned}$$

where  $i$  denotes a firm,  $Exporter_{i,2003}$  indicates if firm  $i$  exports in 2003,  $u_i$  is the error term and the remaining variables are defined as above. These regressions are also separately estimated conditional on expansions and contractions.

**Log-Employment Serial Correlations (Formal Firms)** – see Table M.2

We compute, separately for  $C$ - and  $S$ -sector firms, the serial correlations:

$$Corr(\log \ell_{i,2003}, \log \ell_{i,2004})_k \text{ for } k = C, S$$

**Size Distribution of Formal Firms** – see M.3

We compute, separately for  $C$ - and  $S$ -sector firms, the mean and standard deviation of log-employment across all firms, and the mean of log-employment across all  $C$ -sector exporters.

**Fraction of Exporters** – see Table M.4

We compute the share of all formal  $C$ -sector firms that export.

**Log-Wages (Formal Firms)** – see Table M.5

We compute, separately for  $C$ - and  $S$ -sector firms, the mean of log-wages across all formal firms.

**Log-wage Regressions (Formal Firms)** – see Table M.5

We estimate the following regressions, conditional on  $C$ - and  $S$ -sector firms, respectively (using data for 2003):

$$\begin{aligned} \log(w_i) &= \alpha_C + \beta_C \log(\ell_i) + \gamma_C Exporter_i + u_i \\ \log(w_i) &= \alpha_S + \beta_S \log(\ell_i) + u_i \end{aligned}$$

where  $i$  denote a firm,  $w_i$  is the (average) wage paid by firm  $i$ ,  $u_i$  is the error term and the remaining variables are defined as above.

## L.2 PIA, PAS and PAC (Firm-Level Surveys) and SECEX

*Pesquisa Industrial Anual* (PIA), *Pesquisa Anual de Comércio* (PAC), and *Pesquisa Anual de Serviços* (PAS) are firm-level surveys, covering the formal manufacturing, retail and service sectors, respectively. Conducted by the Brazilian Statistical Agency (IBGE), they contain detailed information on firms' inputs, output and revenues. They constitute a census for larger firms and a representative sample for smaller firms. In the manufacturing sector (PIA), all firms with at least 30 employees are part of the census and are surveyed every year, while firms with 5 to 29 employees are randomly sampled. The PAC (retail sector) and PAS (services) surveys have the same design, but have lower size thresholds for firms to be included in the census: firms with 20 employees or more are part of the census, while firms with up to 19 employees are randomly sampled. Finally, firms in PIA, PAS and PAC are also identified by their 8-digit CNPJ codes. Therefore, we are able to match SECEX with PIA to identify exporters. We use these datasets to obtain the following firm-level variables:

- *Annual gross revenues*
- *Export share*: for firm  $i$ , the share of revenues that comes from exports

$$Export\ Share_i = \frac{Value\ of\ Exports_i}{Revenues_i}$$

Using PIA, PAS, PAC and SECEX, we compute the following moments and auxiliary models.

***Distribution of log-revenues*** – see Table M.6

We compute the mean and standard deviation of log-revenues across all firms in the  $C$  and  $S$  sectors.

***Average Export Share***

Average export share among all exporters, used to recover the value of  $d_F$  conditional on  $\sigma_C$ —see Step 4 in section I.1 for details. We obtain that the average export share among exporters equals 0.264.

***Fraction of Aggregate Revenues in the Formal C-Sector that is Exported*** – see Table M.4

Ratio between total exports and total revenues in the (formal)  $C$  sector.<sup>46</sup>

***Serial Correlation of log-Revenues*** – see Table M.6

$\text{Corr}(\log \text{Revenues}_{i,2004}, \log \text{Revenues}_{i,2003})$  separately for the  $C$  and  $S$  sectors. These moments are computed conditional on firms with at least 30 employees for PIA, and conditional on firms with at least 20 employees for PAS and PAC, so that they are part of the census and therefore surveyed in both years.

***Log-Revenues Regressions*** – see Table M.6

We estimate the following regressions, conditional on  $C$ - and  $S$  sector firms (data from 2003):

$$\begin{aligned}\log(\text{Revenues}_i) &= \alpha_C + \beta_C \log(\ell_i) + \gamma_C \text{Exporter}_i + u_i \\ \log(\text{Revenues}_i) &= \alpha_S + \beta_S \log(\ell_i) + u_i\end{aligned}$$

where  $i$  denotes a firm,  $u_i$  is the error term and the remaining variables are defined as above.

### L.3 ECINF (*Pesquisa de Economia Informal Urbana*)

ECINF was collected by IBGE in 1997 and 2003, and was designed to be representative of the universe of *urban* firms with up to five employees (both formal and informal). It is a matched employer-employee dataset that contains information on entrepreneurs, their businesses and employees. We use the same filters for industries we described above. Although a few firms in the dataset have more than five employees, we restrict attention to those with five employees or less so that our sample is consistent with the population the survey targets. We define as informal firms those that do not have a tax registration number, which means that they are not formally registered as a firm.

ECINF is comprised of two main files. The first contains information on businesses (these are small businesses, so there are no multi-establishment firms and we can use firm and establishment interchangeably) and the second contains information on workers. Before merging these data sources, we drop workers who are younger than 18 and older than 64 years old from the individual level data (only 890 observations are dropped). We then aggregate these data up to the firm level, providing us with information on firms' size and wage bill.<sup>47</sup> We merge this information with the first (firm-level) file using a unique firm identifier. Finally, we trim observations below the first percentile of the revenue distribution, which amounts to dropping firms with revenues very close to zero. We generate the following firm-level variables with ECINF:

<sup>46</sup>The denominator comes from PIA's publication, Table 1.5 (pdf included in the replication folder). The two values used to compute the denominator correspond to the entries "Gross Revenues" and "Other Operational Revenues" of manufacturing firms (*Indústria da Transformação*).

<sup>47</sup>Thus, if a firm has employees older than 64 or younger than 18 years old they are not accounted for when we compute firm size.

- *Informality Indicator*: Dummy variable that equals one if the firm is not registered with the tax authorities.
- *Annual gross revenues*
- *Total number of employees*
- *Average wage*: firm's annual wage bill divided by number of people working at the firm. The wage bill includes the self-reported take-home earnings of the owner. For one-person firms, this is equal to the owner's take-home remuneration.

Using ECINF, we compute the following moments and auxiliary models.

***Size Distribution (Informal Firms)*** – see Table [M.7](#)

We compute the following moments of firm-level log-employment separately for *C*- and *S*-sector informal firms: mean and standard deviation.

***Distribution of Revenues (Informal Firms)*** – see Table [M.7](#)

We compute the mean of firms' log-revenues separately for *C*- and *S*-sector informal firms.

***Log-Wages (Informal Firms)*** – see Table [M.7](#)

We compute the mean of firm-level log-wages separately for *C*- and *S*-sector firms.

***Regression of Informal Status Indicator vs. Number of Employees*** – see Table [M.7](#)

$$Informal_i = \alpha_k + \beta_k \ell_i + u_i$$

where  $i$  denotes firms,  $k = C, S$  denotes sector, and  $u_i$  is the error term.

## L.4 PME

We use the *Pesquisa Mensal de Emprego* (PME) survey to obtain information on worker allocations and labor market flows. It is a rotating panel, in which individuals in a given household are interviewed for 4 consecutive months, they rest for 8 months and are then re-interviewed for additional 4 consecutive months, which implies a maximum panel length of 16 months. As in the firm-level data, we exclude those who are employed in the public sector, agriculture, mining, coal, oil and gas industries. As with ECINF, we only keep individuals who are between 18 and 64 years old. In addition, we exclude individuals who are out of the labor force, non-wage (unpaid) employees or employers. Finally, we restrict our attention to the years of 2003 and 2004. Thus, there are three possible states in our sample:

- (i) Formal workers: those who have a formal labor contract, which in Brazil is defined by having a booklet (*carteira de trabalho*) that has been signed by her employer and that registers workers' entire employment history in the formal sector
- (ii) Informal workers: those who do not have a signed booklet (without a formal contract), which includes self-employed workers
- (iii) Unemployed: those who are not employed, but are actively searching for a job

We employ PME to generate the following moments:

**Transition Matrix** – see Table M.1

To obtain the annual transition matrix between states, we first estimate the 4-month transition matrix using information from the first and fourth interviews. Denote this 4-month transition matrix by  $M$ . We then estimate the annual transition matrix by computing  $M^3$ . This is preferable to using information from the first and fifth interviews—which are 12 months apart—given the high attrition rates between the fourth and fifth interviews, which are 8 months apart. This high attrition is common in panel surveys that have similar designs, as the survey unit is a particular address (e.g. an apartment) and individuals may move in and out during the 8-month rest period.

**Workers’ Allocations** – see Table M.1

We use PME’s sample weights to obtain the total number and the shares of individuals in each of the possible labor market statuses: (i) formal worker in the  $C$  sector; (ii) informal worker in the  $C$  sector; (iii) formal worker in the  $S$  sector; (iv) informal worker in the  $S$  sector; and (v) unemployment.

## L.5 IBGE National Accounts

We employ information available from IBGE’s 2000 and 2005 National Accounts to compute the share of final expenditures on sector  $C$  goods,  $\zeta$ , sector  $k$ ’s fraction of intermediate expenditures on sector  $C$  goods,  $\lambda_k$ , and statistics relevant for the estimation of  $\delta_k$ , which drives the importance of labor in sector  $k$ ’s production.

We compute  $\zeta$  using final demand information, excluding Agriculture and Mining to be consistent with the filters we implemented in the datasets above. We obtain  $\zeta = 0.296$ , as is reported in Table 6.

To obtain information on  $\delta_k$  (conditional on  $\sigma_k$ ), we compute:

$$\begin{aligned} \frac{\text{Total Expenditures with Intermediates}_C}{\text{Total Gross Revenues}_C} &= 0.596, \\ \frac{\text{Total Expenditures with Intermediates}_S}{\text{Total Gross Revenues}_S} &= 0.320. \end{aligned}$$

See Step 3 of section I.1 for details on how to use these statistics to obtain  $\delta_C$  and  $\delta_S$ .

Finally, we compute  $\lambda_k$ , for  $k = C, S$ , as:

$$\lambda_k = \frac{\text{Total Expenditure with Sector } C \text{ Intermediates}}{\text{Total Expenditure with Intermediates (across } C \text{ and } S)},$$

leading to  $\lambda_C = 0.645$  and  $\lambda_S = 0.291$ , as is reported in Table 6.

## M Model Fit: Moments Generated by the Model vs. Data

This section compares the moments generated by the model, using our estimates, with those computed from the data. Tables M.1 through M.7 shows that our model is able to replicate several salient features of the data.

Table M.1: Employment Shares and Transition Rates from Unemployment

Moment	Dataset	Model	Data
Share of Employment $Ci$	PME	0.062	0.059
Share of Employment $Cf$	PME	0.097	0.106
Share of Employment $Si$	PME	0.351	0.351
Share of Employment $Sf$	PME	0.303	0.334
Share Unemployment	PME	0.186	0.150
Share Informal Workers (Conditional on Working)	PME	0.508	0.482
Trans. Rate from Unemp. to $Ci$	PME	0.064	0.064
Trans. Rate from Unemp. to $Cf$	PME	0.060	0.050
Trans. Rate from Unemp. to $Si$	PME	0.380	0.389
Trans. Rate from Unemp. to $Sf$	PME	0.161	0.161
Trans. Rate from Unemp. to Unemp.	PME	0.336	0.336
Ratio Trans. to Informal job / Trans. to Formal job	PME	2.013	2.146

Table M.2: Turnover-Related Moments and Auxiliary Models (Formal Sectors)

	Dataset	$C$ sector		$S$ sector	
		Model	Data	Model	Data
Exit Rate	RAIS	0.098	0.103	0.100	0.125
Average Firm-level Turnover	RAIS	0.231	0.505	0.190	0.525
$Corr(\log \ell_{i,t+1}, \log \ell_{i,t})$	RAIS	0.951	0.929	0.938	0.914
$Exit_i = \alpha + \beta \log(\ell_i)$					
Intercept	RAIS	0.153	0.188	0.145	0.185
$\log(\ell_i)$	RAIS	-0.024	-0.045	-0.035	-0.049
$Turnover_i = \alpha + \beta \log(\ell_i) + \gamma Exporter_i$					
Intercept	RAIS+SECEX	0.426	0.741	0.326	0.645
$\log(\ell_i)$	RAIS+SECEX	-0.089	-0.126	-0.106	-0.096
$Exporter_i$	RAIS+SECEX	0.072	0.071		
$Turnover_i = \alpha + \beta \log(\ell_i) + \gamma Exporter_i$ , Conditional on Expansions					
Intercept	RAIS+SECEX	0.379	0.692	0.278	0.690
$\log(\ell_i)$	RAIS+SECEX	-0.094	-0.138	-0.100	-0.150
$Exporter_i$	RAIS+SECEX	0.118	0.116		
$Turnover_i = \alpha + \beta \log(\ell_i) + \gamma Exporter_i$ , Conditional on Contractions					
Intercept	RAIS+SECEX	0.466	0.744	0.368	0.624
$\log(\ell_i)$	RAIS+SECEX	-0.072	-0.101	-0.067	-0.064
$Exporter_i$	RAIS+SECEX	0.056	0.056		

Table M.3: Firm-Size Distribution (Formal Sectors)

	Dataset	$C$ sector		$S$ sector	
		Model	Data	Model	Data
Avg. Firm-Level log-Employment	RAIS	2.286	1.918	1.287	1.237
Std Dev. Firm-Level log-Employment	RAIS	0.889	1.416	0.667	1.175
Avg. Exporter log-Employment	RAIS + SECEX	3.621	4.014		

Table M.4: Trade-Related Moments

	Dataset	Model	Data
Fraction of Exporters (among formal $C$ -sector firms)	RAIS + SECEX	0.124	0.073
Total Exports / (Total Formal Manufacturing Revenue)	SECEX + IBGE	0.134	0.134

Table M.5: Formal-Sector Wages

		$C$ sector		$S$ sector	
	Dataset	Model	Data	Model	Data
Avg. log-Wages	RAIS	8.630	8.769	8.568	8.567
$\log(w_i) = \alpha + \beta \log(\ell_i) + \gamma \text{Exporter}_i$					
Intercept	RAIS+SECEX	8.298	8.509	8.433	8.436
$\log(\ell_i)$	RAIS+SECEX	0.116	0.117	0.105	0.105
$\text{Exporter}_i$	RAIS+SECEX	0.549	0.462		

Table M.6: Formal-Sector Revenues

		$C$ sector		$S$ sector	
	Dataset	Model	Data	Model	Data
Avg. log-Revenues	IBGE	12.682	12.726	11.151	10.814
Std. Dev. log-Revenues	IBGE	1.241	1.874	0.966	1.440
$\text{Corr}(\log \text{Rev}_{i,t+1}, \log \text{Rev}_{i,t})$	IBGE	0.777	0.929	0.756	0.845
$\log(\text{Rev}_i) = \alpha + \beta \log(\ell_i) + \gamma \text{Exporter}_i$					
Intercept	IBGE+SECEX	9.997	10.118	9.671	10.004
$\log(\ell_i)$	IBGE+SECEX	1.144	1.000	1.150	0.872
$\text{Exporter}_i$	IBGE+SECEX	0.564	1.462		

Notes: The serial correlation of  $\log(\text{Rev})$  is conditional on the employment cutoffs the PIA (30 employees) and PAS (20 employees) panels.

Table M.7: Informal Sector Moments and Auxiliary Moments

		$C$ sector		$S$ sector	
	Dataset	Model	Data	Model	Data
Average log-Employment	ECINF	0.249	0.105	0.275	0.097
Std. Dev. log-Employment	ECINF	0.361	0.303	0.373	0.274
Avg. log-Revenue	ECINF	9.827	8.531	9.254	8.953
Avg. log-Wages	ECINF	7.819	8.043	7.603	8.440
$\text{Informal}_i = \alpha + \beta \ell_i$					
Intercept	ECINF	1.328	1.135	1.244	1.130
$\ell_i$	ECINF	-0.179	-0.179	-0.204	-0.204

Notes: All statistics are computed conditional on firms with five employees or less, both in the data and in the model.



## N Background: The Cost of Labor Regulations in Brazil

The relevant laws and regulations that apply to formal labor relations in Brazil are contained in the Brazilian Labor Code (*Consolidação das Leis Trabalhistas*—CLT). According to the employment index in [Botero et al. \(2004\)](#), the cost of labor regulations in Brazil is around 20 percent above the mean and median of 85 countries and more than 2.5 times larger than in the United States.

The main aspects of labor regulations in Brazil regarding their magnitude and potential impacts on the labor market, are: (i) the presence of a national minimum wage; (ii) unemployment insurance that is only available to formal workers; (iii) substantial firing costs; and (iv) sizable payroll taxes. Since these play an important role in our model and counterfactuals, we provide a brief background discussion of each of them individually.

The nominal value of the national minimum wage is determined by the federal government once a year and is typically binding for many firms. For instance, in 2003 (the year we use in our empirical analysis), the minimum wage corresponds to 49 percent of the national average wage and 81.3 percent of the national median wage.<sup>48</sup>

Although the regulations governing unemployment insurance (UI) are intricate, typically, workers are eligible to receive UI benefits for a duration of 4 to 5 months. The amount of the benefit is determined by the worker’s average wage during the three months preceding their layoff. The replacement rate is 100 percent for individuals who earn one minimum wage, with an average replacement rate of 64 percent (all data come from [Gerard and Gonzaga, 2021](#)).<sup>49</sup>

As for the firing costs, the Brazilian labor regulation states that all formal workers with unjustified dismissal should receive a monetary compensation paid by the employer. In Brazil (and most Latin American countries), firms’ outcomes (e.g. lack of businesses) are not considered a just cause for dismissal, thus any involuntary separation falls in this category ([Heckman and Pagés, 2000](#)). The magnitude of this compensation is determined as a fraction of the funds accumulated in the worker’s *Fundo de Garantia por Tempo de Serviço* (FGTS), which is a job security fund proportional to job tenure and accumulates at a rate of roughly one monthly wage per year. Firms hand over additional severance payments to workers and a direct “penalty” to the government, which further increase the magnitude of the firing costs.<sup>50</sup>

Finally, Brazil has a burdensome tax system, which is characterized not only by high tax rates, but also by a complex structure that implies large compliance costs. For instance, the time required to comply with labor taxes in Brazil is almost 5 times higher than in the U.S. (491 and 100 hours, respectively).<sup>51</sup>

## O Additional Post-Estimation Results: Overlapping Distributions of Productivity

This section presents an important post-estimation result. Our model generates overlapping distributions of firm-level total factor productivity across formal and informal firms within sector  $k \in \{C, S\}$ , an empirical fact that has been emphasized in previous work (e.g., [Meghir et al., 2015](#)). Figure O.1 shows this overlap in

---

<sup>48</sup>The mean and the median wages are computed using micro data from the National Household Survey (PNAD) and pooling together all formal and informal employees who are between 18 and 64 years old and work at least 20 hours per week.

<sup>49</sup>We focus on the rules in place before the 2015 reforms since our empirical analysis precedes them.

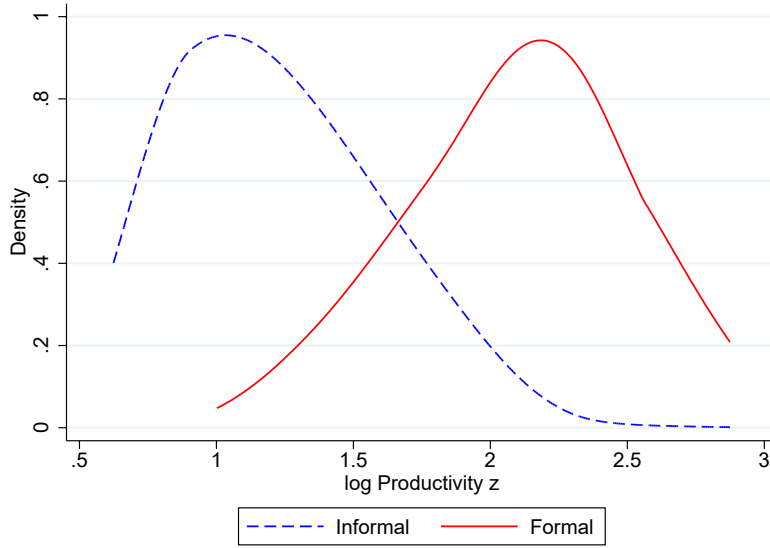
<sup>50</sup>[Gonzaga et al. \(2003\)](#) provide an in depth discussion of the legislation on dismissal costs in Brazil.

<sup>51</sup>These data come from [Doing Business \(2007\)](#), which is the earliest report available on paying taxes in the Doing Business Initiative that provides comparability across a comprehensive set of countries.

sector  $C$ , and Figure O.2 shows this overlap in sector  $S$ .

There are two reasons why the model generates this overlap. First, and more importantly, we do not allow formal firms to switch to informal status. This generates hysteresis: formal firms (which are more productive than informal firms at the entry stage) keep their status even if their productivity significantly declines over time, below the entry cutoff for formal firms. Second, selection into formality/informality depends on a two-dimensional state  $(z, \ell)$ . Therefore, for a fixed value of  $z$ , we can have selection into informality for some levels of  $\ell$  and into formality for other values of  $\ell$ , also contributing to an overlap in the two productivity distributions.

Figure O.1: Kernel Densities of Firm-Level log Productivity— $\log(z)$ —in the Formal and Informal Sectors, sector  $C$



## P Additional Post-Estimation Results: Revenue per Worker is Increasing with Firm-Level Total Factor Productivity

In another post-estimation exercise, we show that revenue per worker within sector  $k \in \{C, S\}$  is informative about firm-level total factor productivity  $z$ . Therefore, from the perspective of our model, Fact 3 can be approximately stated as “informal firms are, on average, less productive than formal firms.”

This post-estimation exercise is important because, in many models, equilibrium revenue per worker is not informative about firm-level productivity  $z$ . For example, in a standard Melitz model, revenue per worker is constant in equilibrium. In our model, revenue per worker varies across firms in equilibrium for many reasons. In particular, as discussed in section 3.3, convex hiring costs imply equilibrium dispersion in revenue per worker. Other reasons contributing to this dispersion include the unequal incidence of taxes and regulations.

Table P.1 shows that our model generates an equilibrium in which revenue per worker increases with productivity. It presents the results of regressions of  $\log\left(\frac{R_k(z, \ell)}{\ell}\right)$  on  $\log(z)$  for  $k \in \{C, S\}$ , using data generated by our model. Firm-level total factor productivity predicts revenue per worker with a high  $R^2$ .

Table P.2 shows a sizable gap between revenue per worker across formal and informal firms. Average

Figure O.2: Kernel Densities of Firm-Level log Productivity— $\log(z)$ —in the Formal and Informal Sectors, sector  $S$

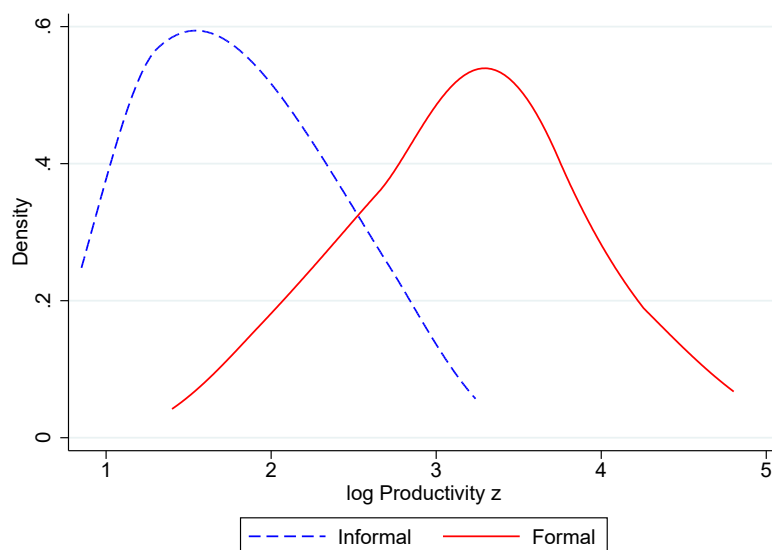


Table P.1: Revenue per Worker Strongly Correlates with Productivity

	Dependent Variable: log Rev. per Worker							
	(1) All $C$	(2) All $C$	(3) Informal $C$	(4) Formal $C$	(5) All $S$	(6) All $S$	(7) Informal $S$	(8) Formal $S$
$\log(z)$	1.321 (0.003)	1.555 (0.002)	1.604 (0.007)	1.274 (0.004)	0.689 (0.001)	0.731 (0.002)	0.729 (0.004)	0.735 (0.003)
<i>Formal</i>		-0.547 (0.003)				-0.125 (0.004)		
Observations	23989	23989	927	23062	16660	16660	1883	14777
$R^2$	0.916	0.964	0.984	0.798	0.932	0.936	0.944	0.809

revenue per worker is substantially larger in the formal sector. There is also a large dispersion in revenue per worker within each sector/firm type combination.

Table P.2: Summary Statistics for Revenue per Worker

	Informal	Formal	Overall
Avg. log Rev per Worker, Sector $C$	8.246	9.065	8.354
Std. Dev. log Rev per Worker, Sector $C$	0.611	0.578	0.667
Avg. log Rev per Worker, Sector $S$	8.336	9.220	8.526
Std. Dev. log Rev per Worker, Sector $S$	0.431	0.596	0.595

## Q Sensitivity Analysis

### Sensitivity Of The Loss Function With Respect To Parameters

In Figures Q.1 to Q.4, we normalize the loss function relative to its optimal value, and we normalize each parameter relative to its estimated value. In each plot, we vary a specified parameter fixing the remaining ones at estimated values. Overall, the figures show that the loss function is sensitive to all the parameters we estimate.

Notice that some figures exhibit small discontinuities with respect to specific parameters. These small discontinuities are explained by the discrete grid we employ for both productivities and employment. As these grids become finer, these discontinuities become smaller and smaller. However, there is a trade-off between finer grids and computational time.

Figure Q.1: Loss Function Sensitivity

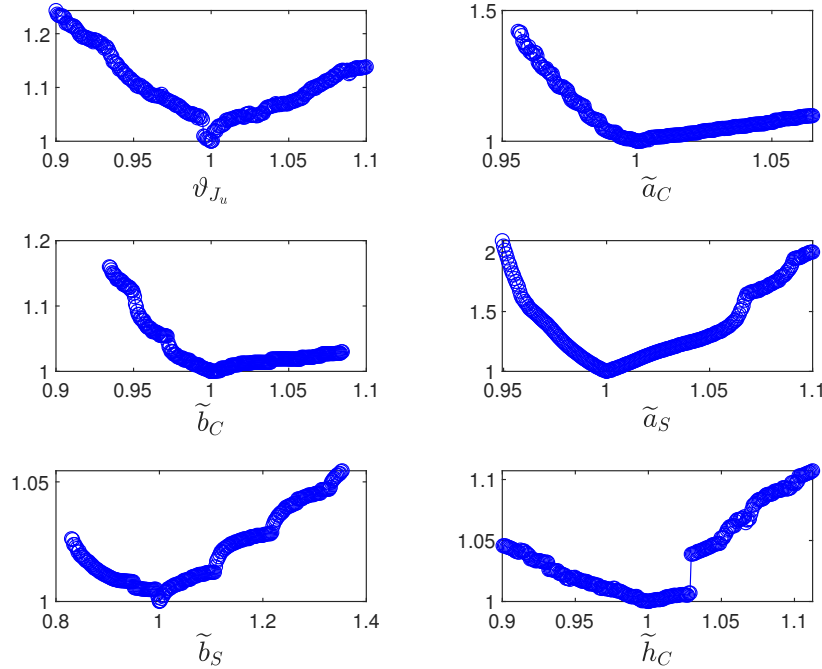


Figure Q.2: Loss Function Sensitivity

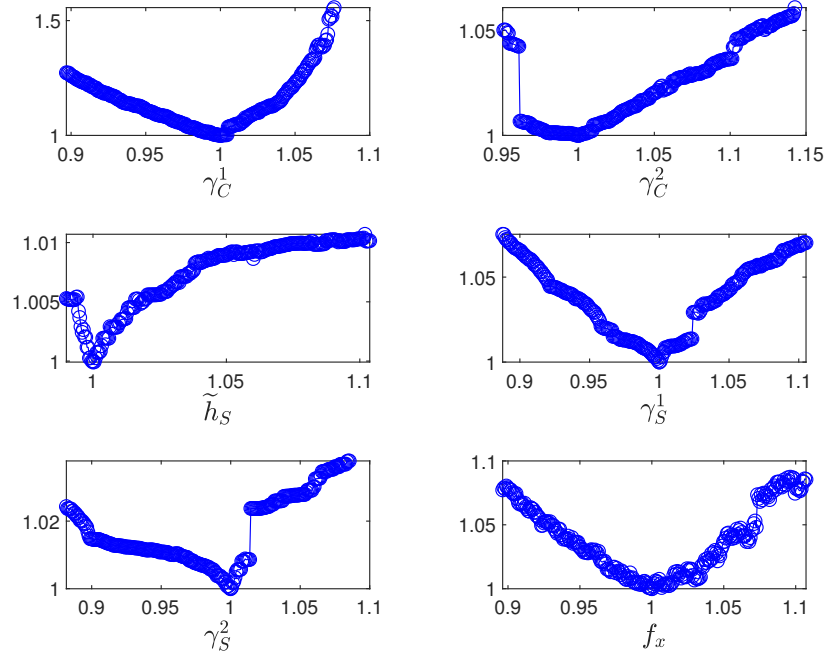


Figure Q.3: Loss Function Sensitivity

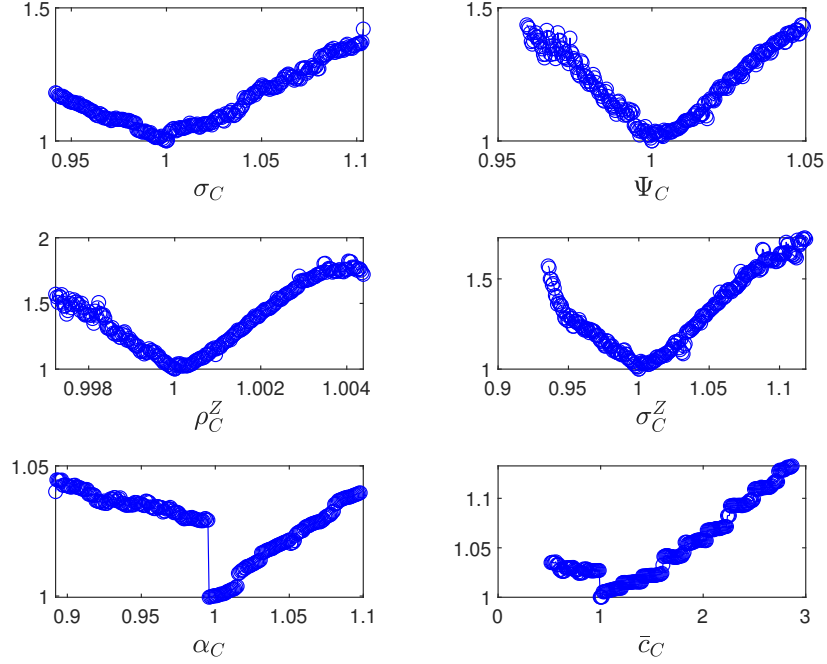
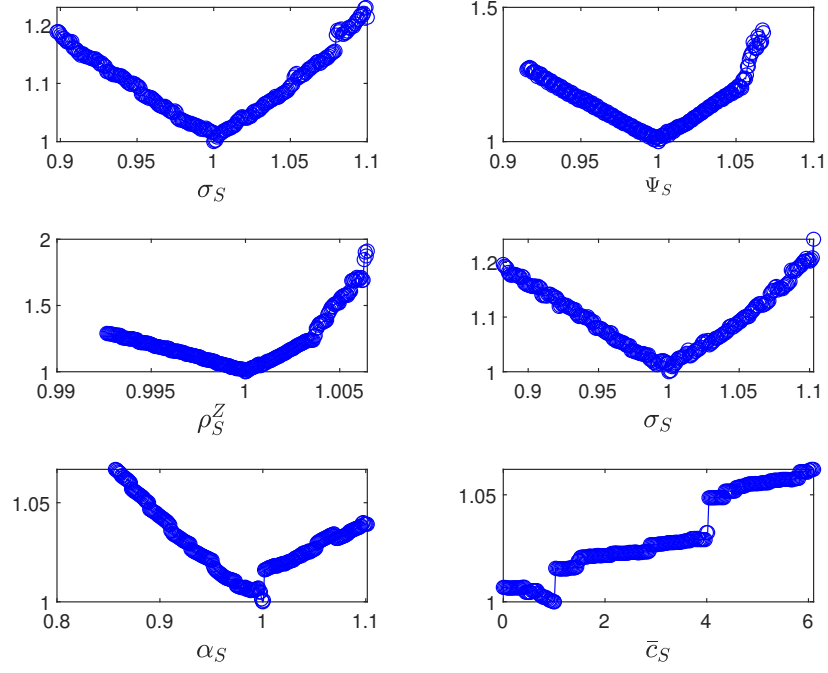


Figure Q.4: Loss Function Sensitivity



## Elasticities Of Moments With Respect To Parameters

Tables Q.1 to Q.6 display the elasticity of each moment used in the loss function relative to each parameter. Here are a few lessons from these tables:

- $\vartheta_{J_u}$  measures the outside option of unemployed workers. In theory, this “parameter” directly affects the level of wages. Indeed, it is shown to affect the levels of  $Cf$ ,  $Ci$  and  $Si$  wages. The impact is less strong on  $Sf$  wages, presumably because of minimum wages that are more binding in that sector. The behavior of the various elasticities with respect to this parameter illustrate the rich interactions in our model: moving this object also has large impacts on various employment shares as well as on turnover moments.
- The parameters determining the expected costs of informality,  $\tilde{a}_k$  and  $\tilde{b}_k$  for  $k \in \{C, S\}$ , strongly affect shares on employment in  $ki$  and they strongly impact the regression relating the probability of informal status to the size of the firm. They also affect wages in  $kf$  and  $ki$ .
- The hiring cost parameters  $\tilde{h}_k$ ,  $\gamma_k^1$ ,  $\gamma_k^2$ ,  $k \in \{C, S\}$ , affect turnover rates and wage regressions.
- The fixed cost of exporting  $f_x$  affects the fraction of exporters.
- The elasticity of substitution  $\sigma_k$  for  $k \in \{C, S\}$ , affects a large number of moments. They strongly affect the fraction of exporters (in the  $C$  sector) and the ratio of exports to revenue in  $Cf$ . Importantly, it affects the log-revenue regressions—both the constant and the slope.
- The persistence of the AR(1) process,  $\rho_k^Z$  for  $k \in \{C, S\}$ , also impacts many moments, but importantly, it strongly affects the serial correlation in revenues at the firm level.
- The standard deviation of the stochastic shocks in the AR(1) process,  $\sigma_k^Z$  for  $k \in \{C, S\}$ , is an important driver of firm size dispersion.
- Exogenous exit rates,  $\alpha_k$  for  $k \in \{C, S\}$ , drive aggregate exit rates.
- Fixed costs of operation  $\bar{c}_k$  for  $k \in \{C, S\}$ , impact the coefficient on size in the exit regressions.

Table Q.1: Elasticity Of Each Moment With Respect To Each Parameter

	$\vartheta_{J_u}$	$\tilde{a}_C$	$\tilde{b}_C$	$\tilde{a}_S$	$\tilde{b}_S$	$\tilde{h}_C$	$\gamma_C^1$	$\gamma_C^2$
1. Share Emp $Ci$	0.102	-2.367	-0.634	0.653	0.011	0.166	0.230	0.021
2. Share Emp $Cf$	0.211	1.470	0.318	0.653	0.011	-0.104	-0.010	0.017
3. Share Emp $Si$	-0.987	0.002	0.012	-7.709	-0.124	0.002	-0.020	-0.005
4. Share Emp $Sf$	1.067	0.002	0.012	9.073	0.139	0.002	-0.020	-0.005
5. Share Emp $U$	-0.023	0.006	-0.000	0.073	-0.002	-0.009	0.001	0.002
6. Tr $U Ci$	0.243	-1.708	-0.555	0.579	0.012	0.128	0.658	-0.015
7. Tr $U Cf$	0.088	1.846	0.477	0.579	0.012	-0.242	-0.501	0.084
8. Tr $U Si$	-0.543	-0.004	0.012	-5.434	-0.074	0.012	-0.021	-0.007
9. Tr $U Sf$	1.155	-0.004	0.012	13.836	0.165	0.012	-0.021	-0.007
10. Tr $U U$	0.000	0.000	-0.000	-0.000	0.000	-0.000	-0.000	0.000
11. Exit Rate $Cf$	0.118	0.187	0.027	0.000	-0.000	-0.064	-0.136	0.039
12. Exit Rate $Cf$ Const	0.079	0.119	-0.019	-0.000	0.000	-0.015	0.335	-0.008
13. Exit Rate $Cf$ Size	-0.137	0.930	0.098	-0.000	0.000	0.254	1.924	-0.212
14. Exit Rate $Sf$	0.155	-0.000	-0.000	1.193	0.020	-0.000	-0.000	0.000
15. Exit Rate $Sf$ Const	0.087	-0.000	-0.000	0.098	0.032	-0.000	0.000	-0.000
16. Exit Rate $Sf$ Size	0.223	-0.000	-0.000	3.355	0.121	-0.000	0.000	0.000
17. exp(mean log(size) $Cf$ )	0.332	-2.131	-0.452	0.000	0.000	-0.414	-1.631	0.272
18. exp(SD log(size) $Cf$ )	-0.417	0.149	0.036	0.000	0.000	-0.024	-0.179	0.032
19. exp(mean log(size)—exp. $Cf$ )	-0.346	0.034	-0.003	0.000	0.000	-0.433	-1.837	0.299
20. exp(mean log(size) $Sf$ )	-0.368	0.000	0.000	-7.315	-0.079	0.000	-0.000	0.000
21. exp(SD log(size) $Sf$ )	-0.162	0.000	0.000	-0.084	-0.015	0.000	-0.000	0.000
22. Fraction Export	0.259	-3.160	-0.610	0.000	0.000	-0.392	-2.023	0.328
23. mean Turnover $Cf$	-0.376	0.419	0.122	-0.000	0.000	-0.107	-0.098	0.068
24. mean Turnover $Sf$	-0.212	0.000	-0.000	2.661	0.026	0.000	-0.000	-0.000
25. Turn. $Cf$ Const	0.021	-0.106	0.007	-0.000	0.000	-0.186	-0.209	0.110
26. Turn. $Cf$ Size	0.398	0.144	0.058	-0.000	0.000	-0.136	0.164	0.078
27. Turn. $Cf$ Exp	1.265	1.155	0.215	0.000	-0.000	-0.702	-2.960	0.666
28. Turn. $Sf$ Const	-0.119	-0.000	-0.000	-0.815	-0.004	-0.000	0.000	-0.000
29. Turn. $Sf$ Size	0.293	-0.000	-0.000	0.085	0.016	-0.000	0.000	-0.000
30. Turn. $Cf$ — expand Const	0.709	-0.382	0.094	0.000	-0.000	-0.436	-1.018	0.108
31. Turn. $Cf$ — expand Size	1.220	-0.486	0.138	0.000	-0.000	-0.461	-1.030	0.029
32. Turn. $Cf$ — expand export	1.103	-0.257	0.095	0.000	-0.000	-0.468	-2.049	0.202
33. Turn. $Sf$ — expand Const	-0.018	-0.000	-0.000	-0.137	-0.029	-0.000	0.000	-0.000
34. Turn. $Sf$ — expand Size	0.086	-0.000	-0.000	0.564	-0.028	-0.000	0.000	0.000
35. Turn. $Cf$ — shrink Const	-0.260	-0.095	-0.028	-0.000	-0.000	-0.066	0.044	0.133
36. Turn. $Cf$ — shrink Size	0.838	0.228	0.019	-0.000	-0.000	-0.213	-0.090	0.324
37. Turn. $Cf$ — shrink Exp	3.402	1.767	0.290	0.000	-0.000	-4.003	-15.717	2.884
38. Turn. $Sf$ — shrink — Const	-0.186	-0.000	-0.000	-1.068	-0.006	-0.000	0.000	-0.000
39. Turn. $Sf$ — shrink Size	3.012	-0.000	-0.000	2.207	0.068	-0.000	0.000	0.000



Table Q.2: Elasticity Of Each Moment With Respect To Each Parameter – Continued

	$\vartheta_{J_u}$	$\tilde{a}_C$	$\tilde{b}_C$	$\tilde{a}_S$	$\tilde{b}_S$	$\tilde{h}_C$	$\gamma_C^1$	$\gamma_C^2$
40. exp(mean log(Wages) <i>Cf</i> )	0.422	-0.591	-0.078	0.000	0.000	-0.019	-0.347	0.027
41. exp(mean log(Wages) <i>Sf</i> )	0.069	0.000	0.000	-2.329	-0.051	-0.000	0.000	0.000
42. exp(log(w) <i>Cf</i> Const)	0.742	-0.750	-0.025	0.000	0.000	-0.126	-0.825	0.152
43. log(w) <i>fC</i> Size	-1.595	2.540	0.175	-0.000	0.000	0.635	2.745	-0.646
44. log(w) <i>fC</i> Exp	0.512	-0.636	-0.090	-0.000	-0.000	0.210	1.392	-0.130
45. exp(log(w) <i>fS</i> Const)	0.085	0.000	0.000	-2.590	-0.093	0.000	-0.000	0.000
46. log(w) <i>fS</i> Size	0.162	-0.000	-0.000	8.392	0.374	-0.000	0.000	0.000
47. <i>Corr(emp<sub>t</sub>, emp<sub>t+1</sub>) Cf</i>	0.012	-0.019	-0.007	0.000	0.000	0.000	-0.022	0.003
48. <i>Corr(emp<sub>t</sub>, emp<sub>t+1</sub>) Sf</i>	0.038	0.000	-0.000	-0.075	-0.004	0.000	-0.000	0.000
49. exp(mean log(size) <i>Ci</i> )	-0.101	-0.709	-0.220	-0.000	-0.000	-0.033	-0.063	-0.003
50. exp(SD log(size) <i>Ci</i> )	-0.064	-0.643	-0.238	-0.000	0.000	-0.031	-0.057	0.012
51. exp(mean log(size) <i>Si</i> )	-0.369	-0.000	0.000	-2.156	-0.046	0.000	-0.000	0.000
52. exp(SD log(size) <i>Si</i> )	-0.302	-0.000	0.000	-2.012	-0.049	0.000	-0.000	-0.000
53. exp(mean log(rev) <i>Ci</i> )	0.637	-0.678	-0.106	-0.000	-0.000	-0.129	-0.430	-0.005
54. exp(mean log(rev) <i>Si</i> )	0.324	-0.000	0.000	-3.445	-0.065	0.000	-0.000	-0.000
55. exp(mean log(Wages) <i>Ci</i> )	0.818	-0.578	0.093	-0.000	0.000	-0.071	-0.275	0.015
56. exp(mean log(Wages) <i>Si</i> )	0.821	0.000	0.000	-1.826	-0.027	-0.000	0.000	-0.000
57. Inf Dummy <i>C</i> Const	0.170	2.664	1.118	-0.000	-0.000	0.060	0.309	-0.079
58. Inf Dummy <i>C</i> Slope	0.561	13.721	5.016	-0.000	-0.000	0.278	1.403	-0.345
59. Inf Dummy <i>S</i> Const	0.243	0.000	0.000	0.337	0.065	-0.000	0.000	-0.000
60. Inf Dummy <i>S</i> Slope	1.740	0.000	-0.000	14.248	0.327	-0.000	0.000	-0.000
61. exp(mean log(rev) <i>Cf</i> )	0.695	-2.776	-0.538	0.000	0.000	-0.400	-1.892	0.285
62. exp(SD log(rev) <i>Cf</i> )	-0.413	0.070	0.011	0.000	-0.000	0.038	0.034	-0.021
63. exp(mean log(rev) <i>Sf</i> )	-0.414	0.000	0.000	-9.982	-0.137	0.000	-0.000	-0.000
64. exp(SD log(rev) <i>Sf</i> )	-0.177	0.000	0.000	-0.276	-0.005	0.000	-0.000	-0.000
65. ratio Exports Revenue <i>Cf</i>	-0.078	-0.928	-0.178	0.000	0.000	-0.209	-1.054	0.182
66. <i>Corr(rev<sub>t</sub>, rev<sub>t+1</sub>) Cf</i>	0.033	-0.000	-0.000	0.000	0.000	-0.041	-0.167	0.023
67. <i>Corr(rev<sub>t</sub>, rev<sub>t+1</sub>) Sf</i>	0.062	0.000	-0.000	0.000	0.000	0.000	-0.000	-0.000
68. exp(log(rev) Const <i>Cf</i> )	0.713	-0.800	-0.023	0.000	0.000	-0.130	-0.857	0.171
69. log(rev) log(size) <i>Cf</i>	-0.179	0.281	0.020	-0.000	-0.000	0.083	0.333	-0.082
70. log(rev) Exp <i>Cf</i>	0.719	-0.738	-0.107	-0.000	-0.000	0.192	1.391	-0.110
71. exp(log(rev) Const <i>Sf</i> )	0.019	0.000	0.000	-2.756	-0.106	0.000	-0.000	0.000
72. log(rev) log(size) <i>Sf</i>	-0.008	-0.000	-0.000	0.991	0.041	-0.000	0.000	0.000
73. Tr. Informal / Tr. Formal	-1.296	-0.742	-0.206	-14.234	-0.186	0.085	0.227	-0.026
74. Share Informal Workers	-0.829	-0.349	-0.084	-6.385	-0.105	0.025	0.017	-0.001

Table Q.3: Elasticity Of Each Moment With Respect To Each Parameter – Continued

	$\tilde{h}_S$	$\gamma_S^1$	$\gamma_S^2$	$f_x$	$\sigma_C$	$\Psi_C$	$\rho_C^Z$	$\sigma_C^Z$
1. Share Emp $Ci$	0.014	-0.082	-0.013	0.072	-4.051	-8.040	-153.831	-6.451
2. Share Emp $Cf$	0.014	-0.082	-0.013	0.039	1.197	0.804	15.022	0.467
3. Share Emp $Si$	-0.231	0.309	0.141	-0.013	0.207	0.692	13.116	0.569
4. Share Emp $Sf$	0.266	-0.357	-0.159	-0.013	0.207	0.692	13.116	0.569
5. Share Emp $U$	-0.010	0.069	0.004	0.000	0.001	-0.129	-2.439	-0.015
6. Tr $U Ci$	0.024	-0.151	-0.017	0.085	-3.325	-7.845	-147.555	-5.832
7. Tr $U Cf$	0.024	-0.151	-0.017	0.026	1.710	1.133	17.990	1.188
8. Tr $U Si$	-0.112	0.289	0.102	-0.013	0.206	0.821	15.556	0.584
9. Tr $U Sf$	0.244	-0.568	-0.230	-0.013	0.206	0.821	15.556	0.584
10. Tr $U U$	-0.000	0.000	-0.000	0.000	0.000	-0.000	0.000	-0.000
11. Exit Rate $Cf$	-0.000	0.000	0.000	-0.006	0.394	-0.874	-16.612	-0.101
12. Exit Rate $Cf$ Const	0.000	-0.000	-0.000	0.149	-0.596	-2.860	-50.858	-1.521
13. Exit Rate $Cf$ Size	0.000	-0.000	-0.000	0.452	-3.306	-8.002	-145.822	-5.249
14. Exit Rate $Sf$	0.048	-0.036	-0.001	-0.000	-0.000	-0.000	0.000	0.000
15. Exit Rate $Sf$ Const	0.092	0.071	0.061	0.000	0.000	-0.000	0.000	-0.000
16. Exit Rate $Sf$ Size	0.419	-0.260	-0.046	0.000	0.000	-0.000	0.000	0.000
17. exp(mean log(size) $Cf$ )	-0.000	0.000	0.000	-0.054	2.102	3.623	74.681	2.754
18. exp(SD log(size) $Cf$ )	-0.000	-0.000	-0.000	-0.095	1.271	3.006	51.687	2.308
19. exp(mean log(size)—exp. $Cf$ )	-0.000	-0.000	0.000	0.184	3.409	5.720	98.773	3.607
20. exp(mean log(size) $Sf$ )	-0.295	0.729	0.310	-0.000	-0.000	0.000	0.000	0.000
21. exp(SD log(size) $Sf$ )	-0.093	0.106	0.048	-0.000	-0.000	0.000	0.000	0.000
22. Fraction Export	-0.000	0.000	-0.000	-1.629	4.082	13.466	260.999	12.628
23. mean Turnover $Cf$	0.000	-0.000	0.000	-0.007	1.201	0.620	6.414	1.274
24. mean Turnover $Sf$	0.007	0.094	0.042	0.000	-0.000	-0.000	0.000	-0.000
25. Turn. $Cf$ Const	0.000	0.000	0.000	0.046	0.696	-0.411	-5.645	0.236
26. Turn. $Cf$ Size	0.000	-0.000	-0.000	0.030	-0.768	-2.513	-39.260	-1.731
27. Turn. $Cf$ Exp	-0.000	0.000	-0.000	-0.629	-2.276	4.612	79.244	0.154
28. Turn. $Sf$ Const	-0.052	0.381	0.187	0.000	0.000	-0.000	0.000	-0.000
29. Turn. $Sf$ Size	0.095	0.209	0.145	0.000	0.000	-0.000	0.000	-0.000
30. Turn. $Cf$ — expand Const	-0.000	-0.000	-0.000	-0.174	0.232	0.125	10.367	0.571
31. Turn. $Cf$ — expand Size	-0.000	-0.000	-0.000	-0.444	-0.420	-0.271	6.601	0.183
32. Turn. $Cf$ — expand export	-0.000	-0.000	-0.000	-0.781	0.785	5.992	105.825	3.924
33. Turn. $Sf$ — expand Const	-0.167	0.565	0.072	0.000	0.000	-0.000	0.000	-0.000
34. Turn. $Sf$ — expand Size	-0.086	0.455	-0.081	0.000	0.000	-0.000	-0.000	-0.000
35. Turn. $Cf$ — shrink Const	0.000	0.000	0.000	0.117	1.323	-0.002	-0.603	0.622
36. Turn. $Cf$ — shrink Size	0.000	0.000	0.000	0.282	-2.643	-5.949	-95.729	-4.649
37. Turn. $Cf$ — shrink Exp	-0.001	0.000	-0.000	-3.027	-3.229	20.274	384.999	4.395
38. Turn. $Sf$ — shrink — Const	-0.026	0.243	0.332	0.000	0.000	-0.000	0.000	-0.000
39. Turn. $Sf$ — shrink Size	0.138	0.513	0.902	0.000	0.000	-0.000	0.000	-0.000

Table Q.4: Elasticity Of Each Moment With Respect To Each Parameter – Continued

	$\tilde{h}_S$	$\gamma_S^1$	$\gamma_S^2$	$f_x$	$\sigma_C$	$\Psi_C$	$\rho_C^Z$	$\sigma_C^Z$
40. exp(mean log(Wages) $Cf$ )	0.000	-0.000	0.000	-0.055	0.114	0.633	16.942	0.797
41. exp(mean log(Wages) $Sf$ )	-0.108	0.097	0.080	0.000	0.000	0.000	-0.000	-0.000
42. exp(log(w) $Cf$ Const)	-0.000	-0.000	0.000	-0.315	0.488	1.111	30.365	1.355
43. log(w) $fC$ Size	0.000	-0.000	0.000	1.425	-3.267	-6.323	-144.652	-5.891
44. log(w) $fC$ Exp	0.000	0.000	-0.000	0.090	0.147	2.429	45.703	2.733
45. exp(log(w) $fS$ Const)	-0.216	0.298	0.224	-0.000	-0.000	0.000	-0.000	-0.000
46. log(w) $fS$ Size	1.031	-2.074	-1.332	0.000	0.000	-0.000	0.000	-0.000
47. $Corr(emp_t, emp_{t+1}) Cf$	-0.000	-0.000	-0.000	-0.012	0.067	0.214	4.054	0.121
48. $Corr(emp_t, emp_{t+1}) Sf$	-0.011	-0.004	0.004	-0.000	-0.000	-0.000	0.000	-0.000
49. exp(mean log(size) $Ci$ )	-0.000	0.000	-0.000	-0.004	0.343	0.391	7.813	0.216
50. exp(SD log(size) $Ci$ )	-0.000	0.000	-0.000	-0.008	0.290	0.351	7.103	0.246
51. exp(mean log(size) $Si$ )	-0.085	0.192	0.058	-0.000	-0.000	0.000	-0.000	0.000
52. exp(SD log(size) $Si$ )	-0.112	0.216	0.059	-0.000	-0.000	0.000	-0.000	0.000
53. exp(mean log(rev) $Ci$ )	-0.000	-0.000	0.000	-0.028	0.123	1.541	28.745	0.667
54. exp(mean log(rev) $Si$ )	-0.160	0.277	0.118	-0.000	-0.000	0.000	0.000	0.000
55. exp(mean log(Wages) $Ci$ )	0.000	0.000	-0.000	-0.020	-0.066	1.051	18.380	0.549
56. exp(mean log(Wages) $Si$ )	-0.053	0.066	0.046	-0.000	0.000	0.000	0.000	-0.000
57. Inf Dummy $C$ Const	0.000	0.000	0.000	0.126	-1.117	-2.367	-41.717	-1.628
58. Inf Dummy $C$ Slope	0.000	-0.000	-0.000	0.482	-4.579	-9.795	-186.380	-7.071
59. Inf Dummy $S$ Const	0.099	-0.097	-0.010	0.000	0.000	0.000	-0.000	0.000
60. Inf Dummy $S$ Slope	0.690	-1.009	-0.326	0.000	0.000	-0.000	0.000	0.000
61. exp(mean log(rev) $Cf$ )	-0.000	-0.000	-0.000	-0.112	2.084	4.062	88.045	3.438
62. exp(SD log(rev) $Cf$ )	0.000	0.000	-0.000	-0.081	1.261	3.412	58.724	2.763
63. exp(mean log(rev) $Sf$ )	-0.404	0.811	0.389	-0.000	-0.000	0.000	-0.000	0.000
64. exp(SD log(rev) $Sf$ )	-0.045	0.009	-0.023	-0.000	-0.000	0.000	-0.000	0.000
65. ratio Exports Revenue $Cf$	-0.000	0.000	-0.000	-0.944	3.464	11.481	209.965	10.402
66. $Corr(rev_t, rev_{t+1}) Cf$	-0.000	-0.000	0.000	0.111	0.883	1.868	31.261	0.899
67. $Corr(rev_t, rev_{t+1}) Sf$	0.020	0.014	0.016	-0.000	-0.000	-0.000	0.000	-0.000
68. exp(log(rev) Const $Cf$ )	-0.000	0.000	-0.000	-0.337	0.464	1.219	32.364	1.405
69. log(rev) log(size) $Cf$	0.000	0.000	-0.000	0.151	-0.407	-0.845	-18.230	-0.740
70. log(rev) Exp $Cf$	0.000	0.000	0.000	0.094	-0.145	1.669	34.678	2.263
71. exp(log(rev) Const $Sf$ )	-0.238	0.315	0.261	-0.000	-0.000	0.000	0.000	-0.000
72. log(rev) log(size) $Sf$	0.117	-0.230	-0.153	0.000	0.000	0.000	-0.000	0.000
73. Tr. Informal / Tr. Formal	-0.277	0.680	0.258	0.004	-0.910	-1.347	-24.097	-1.101
74. Share Informal Workers	-0.197	0.266	0.118	0.000	-0.431	-0.662	-12.443	-0.514

Table Q.5: Elasticity Of Each Moment With Respect To Each Parameter – Continued

	$\alpha_C$	$\bar{c}_C$	$\sigma_S$	$\Psi_S$	$\rho_S^Z$	$\sigma_S^Z$	$\alpha_S$	$\bar{c}_S$
1. Share Emp $Ci$	0.160	-0.100	0.725	1.950	31.816	1.359	-0.060	0.005
2. Share Emp $Cf$	-0.182	0.018	0.725	1.950	31.816	1.359	-0.060	0.005
3. Share Emp $Si$	-0.008	0.006	0.056	0.281	-3.153	-0.092	-0.313	-0.018
4. Share Emp $Sf$	-0.008	0.006	-0.427	-0.817	-7.136	-0.509	0.179	0.016
5. Share Emp $U$	0.070	0.002	-0.015	-0.838	-11.079	-0.167	0.348	0.002
6. Tr $U Ci$	0.550	-0.054	0.740	2.788	42.898	1.526	-0.408	0.003
7. Tr $U Cf$	0.117	0.021	0.740	2.788	42.898	1.526	-0.408	0.003
8. Tr $U Si$	-0.077	0.004	0.034	0.039	-1.778	-0.091	-0.129	-0.009
9. Tr $U Sf$	-0.077	0.004	-0.630	-2.256	-31.473	-0.958	0.617	0.019
10. Tr $U U$	-0.000	0.000	-0.000	0.000	-0.000	-0.000	-0.000	-0.000
11. Exit Rate $Cf$	0.748	0.020	0.000	0.000	-0.000	-0.000	-0.000	-0.000
12. Exit Rate $Cf$ Const	0.486	0.040	0.000	0.000	-0.000	-0.000	-0.000	-0.000
13. Exit Rate $Cf$ Size	-0.021	0.064	0.000	0.000	-0.000	-0.000	-0.000	-0.000
14. Exit Rate $Sf$	-0.000	0.000	-0.175	-1.324	-17.443	-0.375	0.856	0.006
15. Exit Rate $Sf$ Const	-0.000	-0.000	-0.163	-1.384	-15.022	-0.292	0.480	0.009
16. Exit Rate $Sf$ Size	-0.000	0.000	-1.853	-4.867	-60.491	-1.866	0.100	0.017
17. exp(mean log(size) $Cf$ )	0.074	0.030	-0.000	-0.000	0.000	0.000	0.000	0.000
18. exp(SD log(size) $Cf$ )	-0.308	-0.026	-0.000	-0.000	0.000	-0.000	0.000	0.000
19. exp(mean log(size)—exp. $Cf$ )	-0.224	0.002	-0.000	-0.000	0.000	-0.000	0.000	0.000
20. exp(mean log(size) $Sf$ )	0.000	-0.000	2.155	4.290	65.605	2.227	-0.568	-0.001
21. exp(SD log(size) $Sf$ )	0.000	-0.000	0.491	0.684	9.552	0.432	-0.207	-0.003
22. Fraction Export	-0.117	0.018	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
23. mean Turnover $Cf$	0.034	0.012	0.000	0.000	-0.000	-0.000	-0.000	-0.000
24. mean Turnover $Sf$	-0.000	0.000	0.603	-1.524	-24.213	0.013	0.190	0.007
25. Turn. $Cf$ Const	0.275	0.059	0.000	0.000	-0.000	-0.000	-0.000	-0.000
26. Turn. $Cf$ Size	0.536	0.102	0.000	0.000	-0.000	-0.000	-0.000	-0.000
27. Turn. $Cf$ Exp	0.743	0.114	-0.000	-0.000	0.000	-0.000	-0.000	-0.000
28. Turn. $Sf$ Const	-0.000	-0.000	1.050	-0.282	-2.637	0.649	0.067	0.014
29. Turn. $Sf$ Size	-0.000	0.000	-0.040	-1.914	-23.875	-0.217	0.338	0.024
30. Turn. $Cf$ — expand Const	0.343	0.020	-0.000	0.000	0.000	-0.000	-0.000	-0.000
31. Turn. $Cf$ — expand Size	0.513	0.032	-0.000	0.000	-0.000	-0.000	-0.000	-0.000
32. Turn. $Cf$ — expand export	0.474	0.031	-0.000	0.000	-0.000	-0.000	-0.000	-0.000
33. Turn. $Sf$ — expand Const	-0.000	-0.000	0.511	0.038	-8.921	0.459	0.056	-0.001
34. Turn. $Sf$ — expand Size	-0.000	0.000	0.040	-1.011	-29.282	-0.001	0.192	-0.001
35. Turn. $Cf$ — shrink Const	0.321	0.087	0.000	0.000	0.000	-0.000	-0.000	-0.000
36. Turn. $Cf$ — shrink Size	0.749	0.184	0.000	-0.000	0.000	0.000	-0.000	-0.000
37. Turn. $Cf$ — shrink Exp	-0.251	0.174	-0.000	0.000	-0.000	-0.000	-0.000	-0.000
38. Turn. $Sf$ — shrink — Const	-0.000	0.000	1.768	-0.275	0.853	1.030	0.203	0.030
39. Turn. $Sf$ — shrink Size	-0.000	0.000	-6.435	-9.017	-123.284	-4.540	1.082	0.079

Table Q.6: Elasticity Of Each Moment With Respect To Each Parameter – Continued

	$\alpha_C$	$\bar{c}_C$	$\sigma_S$	$\Psi_S$	$\rho_S^Z$	$\sigma_S^Z$	$\alpha_S$	$\bar{c}_S$
40. exp(mean log(Wages) $Cf$ )	0.145	0.008	-0.000	-0.000	-0.000	-0.000	0.000	-0.000
41. exp(mean log(Wages) $Sf$ )	0.000	-0.000	0.963	1.850	24.394	1.523	-0.000	-0.000
42. exp(log(w) $Cf$ Const)	0.314	0.034	-0.000	-0.000	-0.000	-0.000	0.000	-0.000
43. log(w) $fC$ Size	-0.708	-0.127	0.000	0.000	-0.000	-0.000	-0.000	-0.000
44. log(w) $fC$ Exp	0.244	0.029	0.000	0.000	0.000	0.000	-0.000	0.000
45. exp(log(w) $fS$ Const)	0.000	-0.000	1.051	1.371	2.407	1.450	0.102	0.004
46. log(w) $fS$ Size	-0.000	0.000	-2.338	-1.232	25.431	-1.260	-0.325	-0.035
47. $Corr(emp_t, emp_{t+1}) Cf$	-0.021	-0.001	-0.000	0.000	-0.000	-0.000	-0.000	-0.000
48. $Corr(emp_t, emp_{t+1}) Sf$	-0.000	0.000	-0.030	0.117	2.129	-0.002	-0.023	-0.000
49. exp(mean log(size) $Ci$ )	0.006	0.034	0.000	-0.000	0.000	0.000	0.000	0.000
50. exp(SD log(size) $Ci$ )	-0.010	0.013	0.000	-0.000	-0.000	0.000	0.000	0.000
51. exp(mean log(size) $Si$ )	0.000	0.000	1.183	1.776	22.937	0.947	-0.159	0.006
52. exp(SD log(size) $Si$ )	0.000	-0.000	0.856	1.291	16.649	0.705	-0.169	0.001
53. exp(mean log(rev) $Ci$ )	0.097	0.135	0.000	-0.000	0.000	0.000	0.000	0.000
54. exp(mean log(rev) $Si$ )	0.000	0.000	1.101	2.756	38.961	1.236	-0.202	0.018
55. exp(mean log(Wages) $Ci$ )	0.042	0.054	0.000	-0.000	-0.000	0.000	0.000	0.000
56. exp(mean log(Wages) $Si$ )	0.000	0.000	-0.148	0.391	8.140	0.105	-0.041	0.005
57. Inf Dummy $C$ Const	0.084	0.003	0.000	0.000	0.000	0.000	0.000	-0.000
58. Inf Dummy $C$ Slope	0.240	0.010	0.000	0.000	-0.000	0.000	-0.000	-0.000
59. Inf Dummy $S$ Const	0.000	-0.000	-0.280	-0.196	-1.824	-0.196	0.090	0.001
60. Inf Dummy $S$ Slope	-0.000	0.000	-3.087	-4.766	-64.977	-2.692	0.879	0.010
61. exp(mean log(rev) $Cf$ )	0.261	0.041	-0.000	-0.000	0.000	-0.000	0.000	0.000
62. exp(SD log(rev) $Cf$ )	-0.348	-0.035	-0.000	-0.000	-0.000	0.000	0.000	0.000
63. exp(mean log(rev) $Sf$ )	0.000	-0.000	3.201	6.238	91.178	3.847	-0.529	0.000
64. exp(SD log(rev) $Sf$ )	0.000	-0.000	0.772	1.421	22.318	1.042	-0.252	-0.007
65. ratio Exports Revenue $Cf$	-0.162	0.003	-0.000	-0.000	0.000	-0.000	-0.000	-0.000
66. $Corr(rev_t, rev_{t+1}) Cf$	-0.027	0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
67. $Corr(rev_t, rev_{t+1}) Sf$	-0.000	-0.000	1.009	2.713	34.152	1.026	-0.014	0.000
68. exp(log(rev) Const $Cf$ )	0.396	0.047	-0.000	-0.000	0.000	0.000	0.000	0.000
69. log(rev) log(size) $Cf$	-0.089	-0.017	0.000	-0.000	0.000	0.000	0.000	0.000
70. log(rev) Exp $Cf$	0.305	0.039	0.000	-0.000	-0.000	-0.000	-0.000	-0.000
71. exp(log(rev) Const $Sf$ )	0.000	-0.000	1.082	1.372	1.287	1.481	0.194	0.008
72. log(rev) log(size) $Sf$	0.000	0.000	-0.264	-0.174	4.255	-0.151	-0.048	-0.005
73. Tr. Informal / Tr. Formal	0.037	-0.013	0.402	1.310	14.921	0.431	-0.509	-0.023
74. Share Informal Workers	0.033	-0.009	0.155	0.347	-0.055	0.093	-0.195	-0.014

# Supplementary Material – Not for Publication

## I Estimation

### I.1 Estimation Algorithm

In this section we describe the estimation algorithm in detail, which we break down into several steps for expositional clarity.

Before we proceed, remember that value added for domestic producers in sector  $k$  is given by:

$$VA_k(z, \ell) = \Theta_k (P_k^m)^{-(1-\delta_k)\Lambda_k} (\exp(d_{H,k}))^{\frac{\sigma_k}{\sigma_k-1}\Lambda_k} (z\ell^{\delta_k})^{\Lambda_k},$$

where

$$P_k^m \equiv \frac{P_C^{\lambda_k} P_S^{1-\lambda_k}}{\lambda_k^{\lambda_k} (1-\lambda_k)^{1-\lambda_k}}, \quad (\text{S.1})$$

$$\Theta_k \equiv \left( \frac{1}{(1-\delta_k)\Lambda_k} \right) \left( \frac{(1-\delta_k)(\sigma_k-1)}{\sigma_k} \right)^{\frac{\sigma_k-1}{\sigma_k-1}\Lambda_k},$$

and

$$\Lambda_k \equiv \frac{\sigma_k-1}{\sigma_k-(1-\delta_k)(\sigma_k-1)}.$$

Rewrite value added for domestic producers as

$$VA_k(z, \ell) = \Theta_k \Psi_k (z\ell^{\delta_k})^{\Lambda_k},$$

with

$$\Psi_k \equiv (P_k^m)^{-(1-\delta_k)\Lambda_k} (\exp(d_{H,k}))^{\frac{\sigma_k}{\sigma_k-1}\Lambda_k}. \quad (\text{S.2})$$

Note that  $\Theta_k$  is a solely a function of model's parameters. On the other hand,  $\Psi_k$  is a function of model's parameters but also of equilibrium objects such as  $P_C$ ,  $P_S$  and  $d_{H,k}$ . In turn, value added for exporters is given by:

$$VA_C(z, \ell) = \Theta_C \Psi_C (\exp(d_F))^{\frac{\sigma_C}{\sigma_C-1}\Lambda_C} (z\ell^{\delta_C})^{\Lambda_C}.$$

It will be convenient to define and work with

$$\vartheta_{J_u} \equiv b_0 \times P_Y + \frac{1}{1+r} J^u.$$

$\Psi_C$ ,  $\Psi_S$ ,  $\vartheta_{J_u}$  are treated as parameters to be estimated along with the remaining ones, but these are all endogenous variables. The procedure below makes sure that the values guessed for  $\Psi_C$  and  $\Psi_S$  are equilibrium outcomes (see Step 9 for details). The number of entrants  $M_C$  and  $M_S$  will be set to match  $\Psi_C$  and  $\Psi_S$ . Given knowledge of  $\vartheta_{J_u}$  and the remaining parameters, we can recover the utility flow of unemployment  $b_0$  and the value of unemployment  $J^u$  post-estimation (as  $P_Y$  is a byproduct of the estimation procedure depending on  $P_C$ ,  $P_S$  and  $\zeta$ ).

**Step 1a:**  $\lambda_C$  and  $\lambda_S$  are obtained from input-output tables and fixed throughout.

**Step 1b:** Fix  $\mu^v$  and obtain  $\phi$  using equation (24):

$$\phi = \left( \frac{\mu^v}{(Transition_{Data}^{U \rightarrow E})^{\frac{\xi-1}{\xi}}} \right)^{\xi}$$

where  $Transition_{Data}^{U \rightarrow E}$  is the transition rate from unemployment to employment in the data.

**Step 2:** Start with a parameter vector guess  $\Omega$ , including values for  $\Psi_C$ ,  $\Psi_S$  and  $\vartheta_{J_u}$ .

**Step 3:** Obtain  $\delta_k$  using  $P_k^m \iota_k(z, \ell) = \frac{(1-\delta_k)(\sigma_k-1)}{\sigma_k} R_k(z, \ell)$ :

$$\begin{aligned} \left( \frac{Total\ Expenditures\ with\ Intermediates_k}{Total\ Gross\ Revenues_k} \right)_{Data} &= \frac{(1-\delta_k)(\sigma_k-1)}{\sigma_k} \\ \Rightarrow \delta_k &= 1 - \frac{\sigma_k}{\sigma_k-1} \left( \frac{Total\ Expenditures\ with\ Intermediates_k}{Total\ Gross\ Revenues_k} \right)_{Data} \\ \left( \frac{Total\ Expenditures\ with\ Intermediates_k}{Total\ Gross\ Revenues_k} \right)_{Data} &\text{ is obtained from input-output tables.} \end{aligned}$$

**Step 4:** Obtain  $d_F$  using equation (33):

$$\begin{aligned} E[Export\ Share|Exporter = 1]_{Data} &= (1 - \exp(-\sigma_C \times d_F)) \\ \Rightarrow d_F &= -\frac{1}{\sigma_C} \log(1 - E[Export\ Share|Exporter = 1]_{Data}) \end{aligned}$$

$E[Export\ Share|Exporter = 1]_{Data}$  is the average share of exporters' gross revenues in sector  $C$  coming from exports, obtained from PIA and SECEX.

**Step 5:** This step solves for wage schedules  $w_{kf}(z, \ell')$ ,  $w_{ki}(z, \ell')$  as well as value functions  $V_{kf}(z, \ell)$ ,  $V_{ki}(z, \ell)$ ,  $J_{kf}^e(z, \ell')$ ,  $J_{ki}^e(z, \ell')$ , and firms' policy functions.

**Step 5a:** Compute value added functions  $VA_k(z, \ell)$ .

**Step 5b:** Compute wage schedules  $w_{kf}(z, \ell')$

- Guess a wage schedule  $w_{kf}(z, \ell')$
- Compute the resulting  $V_{kf}(z, \ell')$  using (14)
- Compute  $J_{kf}^e(z, \ell')$  using (I.1)
- Compute  $w_{kf}^u(z, \ell')$  using equation (28)
- Let  $\hat{w}_{kf}^u(z, \ell') = \omega_0 + \omega_1 \frac{VA_k(z, \ell')}{\ell'}$  be the linear projection of  $w_{kf}^u(z, \ell')$  on  $\left[1, \frac{VA_k(z, \ell')}{\ell'}\right]$
- Update  $w_{kf}(z, \ell') = \max \left\{ \hat{w}_{kf}^u(z, \ell'), b_u + \vartheta_{J_u} - \frac{1}{1+r} J_{kf}^e(z, \ell'), \underline{w} \right\}$
- Restart until convergence

**Step 5c:** Compute wage schedules  $w_{ki}(z, \ell')$

- Guess a wage schedule  $w_{ki}(z, \ell')$
- Compute the resulting  $V_{ki}(z, \ell')$  using (18)
- Compute  $J_{ki}^e(z, \ell')$  using (1.2)
- Compute  $w_{ki}^u(z, \ell')$  using equation (31)
- Let  $\hat{w}_{ki}^u(z, \ell') = \omega_0 + \omega_1 \left(1 - \frac{\sigma_k}{\sigma_k - (1 - \delta_k)(\sigma_k - 1)} p_{ki}(\ell')\right) \frac{V_{ki}(z, \ell')}{\ell'}$  be the linear projection of  $w_{ki}^u(z, \ell')$  on  $\left[1, \left(1 - \frac{\sigma_k}{\sigma_k - (1 - \delta_k)(\sigma_k - 1)} p_{ki}(\ell')\right) \frac{V_{ki}(z, \ell')}{\ell'}\right]$
- Update  $w_{ki}(z, \ell') = \max \left\{ \hat{w}_{ki}^u(z, \ell'), \vartheta_{J_u} - \frac{1}{1+r} J_{ki}^e(z, \ell') \right\}$
- Restart until convergence

**Step 6:** Solve for firms' entry decisions. Compute the fraction of entrants in the formal and informal sectors as follows:

$$\omega_{kf} \equiv \Pr \left( I_k^{formal}(z) = 1 \right) = \int_z I_k^{formal}(z) g_k^e(z) dz$$

$$\omega_{ki} \equiv \Pr \left( I_k^{informal}(z) = 1 \right) = \int_z I_k^{informal}(z) g_k^e(z) dz$$

Therefore, if  $M_k$  is the mass of entrants in sector  $k$ , the masses of formal and informal entrants in sector  $k$  are given by:

$$M_{ki} = \omega_{ki} M_k$$

$$M_{kf} = \omega_{kf} M_k$$

Finally, compute the distribution of  $z$  productivities among entrants, conditional on entry into sector  $kj$ .

$$\psi_{ki}^e(z) = \frac{g_k^e(z) I_k^{informal}(z)}{\int_{\tilde{z}} g_k^e(\tilde{z}) I_k^{informal}(\tilde{z}) d\tilde{z}},$$

$$\psi_{kf}^e(z) = \frac{g_k^e(z) I_k^{formal}(z)}{\int_{\tilde{z}} g_k^e(\tilde{z}) I_k^{formal}(\tilde{z}) d\tilde{z}}.$$

**Step 7:** Compute the steady-state distribution of states. For informal firms, start with a guess for  $\psi_{ki}$ . Then, compute

$$\varrho_{ki}^{exit} = \alpha_k + (1 - \alpha_k) \int_z \int_{\ell} \left( I_{ki}^{exit}(z, \ell) + I_{ki}^{change}(z, \ell) \right) \psi_{ki}(z, \ell) d\ell dz.$$

In steady state  $N_{ki} = (1 - \varrho_{ki}^{exit}) N_{ki} + M_{ki}$ . Therefore, set  $\frac{M_{ki}}{N_{ki}}$ , the fraction of sector  $k$  informal firms that are entrants, to:

$$\boxed{\frac{M_{ki}}{N_{ki}} = \varrho_{ki}^{exit} = \frac{\omega_{ki} M_k}{N_{ki}}}.$$

Now, compute  $\tilde{\psi}_{ki}$ :

$$\tilde{\psi}_{ki}(z, \ell) = \mathcal{I}[\ell = 1] \times \varrho_{ki}^{exit} \times \psi_{ki}^e(z) \\ + \mathcal{I}[\ell \geq 1] \times (1 - \alpha_k) \times \psi_{ki}(z, \ell) I_{ki}^{stay}(z, \ell),$$



and  $\widehat{\psi}_{ki}$ :

$$\widehat{\psi}_{ki}(z, \ell') = \int_{\ell} \widetilde{\psi}_{ki}(z, \ell) \mathcal{I}(L_{ki}(z, \ell) = \ell') d\ell$$

Update  $\psi_{ki}$  with:

$$\psi_{ki}(z', \ell') = \int_z \widehat{\psi}_{ki}(z, \ell') g_k(z'|z) dz,$$

and repeat until convergence of  $\psi_{ki}$ . This converged value of  $\psi_{ki}$  will be used directly in the computation of  $\psi_{kf}$  below.

For formal firms, start with guess for  $\psi_{kf}$  and compute:

$$\begin{aligned} \varrho_{kf}^{exit} &= \alpha_k + (1 - \alpha_k) \int_z \int_{\ell} I_{kf}^{exit}(z, \ell) \psi_{kf}(z, \ell) d\ell dz, \\ \varrho_{ki}^{change} &= (1 - \alpha_k) \int_z \int_{\ell} I_{ki}^{change}(z, \ell) \psi_{ki}(z, \ell) d\ell dz. \end{aligned}$$

In steady state:

$$\begin{aligned} \varrho_{kf}^{exit} N_{kf} &= \varrho_{ki}^{change} \underbrace{N_{ki}}_{\frac{\omega_{ki} M_k}{\varrho_{ki}^{exit}}} + \omega_{kf} M_k \\ &= M_k \left( \frac{\varrho_{ki}^{change}}{\varrho_{ki}^{exit}} \omega_{ki} + \omega_{kf} \right) \end{aligned}$$

So that:

$$\boxed{\frac{M_{kf}}{N_{kf}} = \frac{M_k \omega_{kf}}{N_{kf}} = \frac{\varrho_{kf}^{exit} \omega_{kf}}{\frac{\varrho_{ki}^{change}}{\varrho_{ki}^{exit}} \omega_{ki} + \omega_{kf}}}$$

Also, note that

$$\frac{M_{kf}}{N_{kf}} \times \frac{N_{ki}}{M_{ki}} = \frac{\varrho_{kf}^{exit} \omega_{kf}}{\frac{\varrho_{ki}^{change}}{\varrho_{ki}^{exit}} \omega_{ki} + \omega_{kf}} \frac{1}{\varrho_{ki}^{exit}} = \frac{\varrho_{kf}^{exit} \omega_{kf}}{\varrho_{ki}^{change} \omega_{ki} + \varrho_{ki}^{exit} \omega_{kf}}$$

and

$$\frac{M_{kf}}{N_{kf}} \times \frac{N_{ki}}{M_{ki}} = \frac{\omega_{kf}}{\omega_{ki}} \frac{N_{ki}}{N_{kf}}$$

Therefore,

$$\boxed{\frac{N_{ki}}{N_{kf}} = \frac{\varrho_{kf}^{exit} \omega_{ki}}{\varrho_{ki}^{change} \omega_{ki} + \varrho_{ki}^{exit} \omega_{kf}}}$$

Compute  $\tilde{\psi}_{kf}$  as:

$$\tilde{\psi}_{kf}(z, \ell) = \mathcal{I}[\ell = 1] \times \underbrace{\frac{\varrho_{kf}^{exit} \omega_{kf}}{\varrho_{ki}^{change} \omega_{ki} + \omega_{kf}}}_{\frac{M_{kf}}{N_{kf}}} \psi_{kf}^e(z) + \mathcal{I}[\ell \geq 1] \times \left( (1 - \alpha_k) \psi_{kf}(z, \ell) I_{kf}^{stay}(z, \ell) + (1 - \alpha_k) \underbrace{\frac{\varrho_{kf}^{exit} \omega_{ki}}{\varrho_{ki}^{change} \omega_{ki} + \varrho_{ki}^{exit} \omega_{kf}}}_{\frac{N_{ki}}{N_{kf}}} \psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell) \right)$$

and  $\hat{\psi}_{kf}$  as:

$$\hat{\psi}_{kf}(z, \ell') = \int_{\ell} \tilde{\psi}_{kf}(z, \ell) \mathcal{I}(L_{kf}(z, \ell) = \ell') d\ell.$$

Update  $\psi_{kf}$  with:

$$\psi_{kf}(z', \ell') = \int_z \hat{\psi}_{kf}(z, \ell') g_k(z'|z) dz,$$

and repeat until convergence of  $\psi_{kf}$ .

At this point we have the following objects:  $\psi_{kj}$ ,  $\tilde{\psi}_{kj}$ ,  $\hat{\psi}_{kj}$ ,  $\varrho_{ki}^{exit}$ ,  $\varrho_{ki}^{change}$ ,  $\varrho_{kf}^{exit}$ ,  $\chi_{ki \rightarrow f}^{change}$ ,  $\chi_{kf}^{layoff}$ , and  $\chi_{ki}^{leave}$  (see equations (C.7), (C.11) and (C.13)).

**Step 8:** Obtain the entry costs  $c_{e,k}$  ( $k = C, S$ ):

$$c_{e,k} = V_k^e = \int_z \left[ V_{ki}^e(z) I_k^{informal}(z) + V_{kf}^e(z) I_k^{formal}(z) \right] g_k^e(z) dz$$

These costs will be subtracted from aggregate income, and will be added to the expenditure on  $S$ -sector goods.

**Step 9:** This step solves for masses of entrants  $M_k$ 's, masses of firms  $N_{kj}$ 's, aggregate vacancies  $V_{kj}$ 's and mass of unemployment  $L_u$  consistent with  $\Psi_C$ ,  $\Psi_S$ ,  $d_F$  and  $\mu^v$ .

**Step 9a:** Write aggregate income  $I$  as a function of masses of entrants  $M_C$  and  $M_S$ .

**Step 9b:** Write  $P_C$  and  $P_S$  as functions of  $M_C$  and  $M_S$ .

**Step 9c:** Write  $X_C^{int}$  as a function of  $M_C$  and  $M_S$ .

**Step 9d:** Solve for  $\frac{M_S}{M_C}$  that matches  $\Psi_C$ .

**Step 9e:** Separately pin down  $M_C$  and  $M_S$  using the labor market clearing equation  $\bar{L} - L_u = \sum_{k=C, S, j=i, f} L_{kj}$ .

Express  $M_C$  and  $M_S$  as functions of  $L_u$ .

**Step 9fe:** Express masses of firms  $N_{kj}$  as functions of  $L_u$ .

**Step 9g:** Express aggregate posted vacancies  $V_{kj}$  as functions of  $L_u$ .

**Step 9h:** Use equation for  $\mu^v$  (and the value initially guessed in Step 1 for  $\mu^v$ ) to obtain  $L_u$  consistent with  $\Psi_C$ ,  $\Psi_S$ ,  $d_F$  and  $\mu^v$ .

**Step 9i:** Go back and obtain masses of entrants  $M_k$ 's, masses of firms  $N_{kj}$ 's, and aggregate vacancies  $V_{kj}$ 's.

**Step 9j:** Recover price indices  $P_C$  and  $P_S$ .

**Step 9k:** Compute deviation between government revenues and spending with unemployment insurance  $Dev_T$ .

**Step 10:** Obtain job finding rates  $\mu_{kj}^e$  using aggregate vacancies  $V_{kj}$ 's and mass of unemployment  $L_u$  obtained in Step 9.

$$\mu_{kj}^e = \frac{m_{kj}}{L_u} = \phi \frac{V_{kj}}{\tilde{V}} \left( \frac{\tilde{V}}{L_u} \right)^\xi$$

**Step 11:** Use equations (C.14)-(C.15) to obtain allocations  $L_{Cf}$ ,  $L_{Ci}$ ,  $L_{Sf}$ ,  $L_{Si}$ .

$$\begin{aligned} L_{Ci} &= \frac{\mu_{Ci}^e L_u}{\chi_{Ci}^{leave}} \\ L_{Si} &= \frac{\mu_{Si}^e L_u}{\chi_{Si}^{leave}} \\ L_{Cf} &= \frac{\mu_{Cf}^e L_u + \chi_{Ci \rightarrow f}^{change} L_{Ci}}{\chi_{Cf}^{layoff}} \\ L_{Sf} &= \frac{\mu_{Sf}^e L_u + \chi_{Si \rightarrow f}^{change} L_{Si}}{\chi_{Sf}^{layoff}} \end{aligned}$$

**Step 12:** Compute deviation from the labor market clearing equation:

$$Dev_L = abs \left( \frac{\bar{L} - (L_{Cf} + L_{Ci} + L_{Sf} + L_{Si})}{\bar{L}} \right),$$

**Step 13:** Compute all moments to be matched with those in the data.

**Step 14:** Compute Loss Function. Add Model/Data deviations to equilibrium penalty  $EQ\_Penalty$ . The objective function is therefore given by

$$L = L_{mom} + EQ\_Penalty$$

Where  $L_{mom}$  penalizes deviations between moments in the data and  $EQ\_Penalty$  penalizes deviations from the labor market clearing condition:

$$EQ\_Penalty = W_L Dev_L + W_T abs(\min\{Dev_T, 0\})$$

With  $W_L$  and  $W_T$  denoting large weights and  $Dev_T$  is the relative deviation between government revenues and spending with unemployment insurance (see section 1.2 for details). We highly penalize a negative  $Dev_T$ .

**Step 15:** Optimization routine picks new parameter vector  $\Omega$ . Go back to Step 1 until convergence.

**Step 16 (Post estimation):** Obtain  $J^u$  using

$$\begin{aligned} J^u &= \sum_{k,j} \mu_{kj}^e \int_{\ell} \int_z \bar{J}_{kj}^e(z, L_{kj}(z, \ell)) g_{kj}(z, \ell) dz d\ell \\ &+ \left( 1 - \sum_{k,j} \mu_{kj}^e \right) \vartheta_{J_u}. \end{aligned}$$

**Step 17 (Post estimation):** At this point, we know  $J^u$  and can compute

$$b_0 \times P_Y = \vartheta_{J^u} - \frac{1}{1+r} J^u,$$

**Step 18 (Post-estimation):** Obtain  $D_F^*$  (this is the parameter that we need for the counterfactuals as  $d_F$  is endogenous):

$$D_F^* = \frac{(\exp(\sigma_C \times d_F) - 1) (P_C^m)^{(1-\delta_C)(\sigma_C-1)} \Psi_C^{\frac{\sigma_C-1}{\Lambda_C}}}{\bar{\epsilon}^{\sigma_C} \tau_C^{1-\sigma_C}},$$

where  $\bar{\epsilon}$  is the exchange rate value that balances trade:

$$\bar{\epsilon} = \frac{1}{\tau_a \tau_c} (P_C^m)^{(1-\delta_C)} \Psi_C^{\frac{1}{\Lambda_C}} (\tau_a Exports)^{\frac{1}{1-\sigma_C}}.$$

## I.2 Estimation Algorithm – Further Details

This section details the steps within Step 9 of the estimation procedure.

**Step 9:** This step solves for masses of entrants  $M_k$ 's, masses of firms  $N_{kj}$ 's, aggregate vacancies  $V_{kj}$ 's consistent with  $\Psi_C$ ,  $\Psi_S$  and  $d_F$ .

We start with some definitions... Averages “per firm”. All these quantities can be computed after Step 8, that is, after solving for the steady state distribution of states.

$$Avg\_wbill_{ki} = \int_z \int_{\ell'} [w_{ki}(z, \ell') \ell'] \hat{\psi}_{ki}(z, \ell') d\ell' dz \text{ for } k = C, S$$

$$Avg\_wbill_{kf} = \int_z \int_{\ell'} [\max\{w_{kf}(z, \ell'), \underline{w}\} \ell'] \hat{\psi}_{kf}(z, \ell') d\ell' dz \text{ for } k = C, S$$

$$Avg\_Firing\_Costs_{kf} = \kappa \int_z \int_{\ell} [(\ell - L_{kf}(z, \ell)) (1 - I_{kf}^{hire}(z, \ell))] \tilde{\psi}_{kf}(z, \ell) d\ell dz \text{ for } k = C, S$$

$$Avg\_Hiring\_Costs_{kj} = \int_z \int_{\ell} [H_{kj}(\ell, L_{kj}(z, \ell)) I_{kj}^{hire}(z, \ell)] \tilde{\psi}_{kj}(z, \ell) d\ell dz \text{ for } k = C, S; j = i, f$$

$$Avg\_Revenue_{kj} = \int_z \int_{\ell'} R_k(z, \ell') \hat{\psi}_{kj}(z, \ell') d\ell' dz \text{ for } k = C, S; j = i, f$$

$$Avg\_InfPenalty_{ki} = \int_z \int_{\ell'} [p_{ki}(\ell') R_k(z, \ell')] \hat{\psi}_{ki}(z, \ell') d\ell' dz \text{ for } k = C, S$$

$$Avg\_Vacancies_{kj} = \int_z \int_{\ell} v_{kj}(z, \ell) \tilde{\psi}_{kj}(z, \ell) d\ell dz \text{ for } k = C, S; j = i, f$$

$$Avg\_Exports_{Cf} = (1 - \exp(-\sigma_C \times d_F)) \int_z \int_{\ell'} [R_C(z, \ell') \mathcal{I}_C^x(z, \ell')] \hat{\psi}_{Cf}(z, \ell') d\ell' dz$$

$$Fraction\_Export_{Cf} = \int_z \int_{\ell'} \mathcal{I}_C^x(z, \ell') \hat{\psi}_{Cf}(z, \ell') d\ell' dz$$

$$Avg\_size_{kj} = \int_z \int_{\ell} \ell \psi_{kj}(z, \ell) d\ell dz \text{ for } k = C, S; j = i, f$$

Now, define

$$\begin{aligned} Avg\_Price_{kj} &= \int_z \int_{\ell'} p_{kj}(z, \ell')^{1-\sigma_k} \widehat{\psi}_{kj}(z, \ell') d\ell' dz \\ &= \int_z \int_{\ell'} \left( \frac{R_k(z, \ell')}{q_k(z, \ell', \iota_k(z, \ell'))} \right)^{1-\sigma_k} \widehat{\psi}_{kj}(z, \ell') d\ell' dz \text{ for } k = C, S; j = i, f. \end{aligned}$$

We cannot compute  $Avg\_Price_{kj}$ —given  $\Omega$ ,  $\Psi_C$  and  $\Psi_S$ . However, note that:

$$\begin{aligned} Avg\_Price_{kj} &= \left( \frac{(1-\delta_k)(\sigma_k-1)}{\sigma_k} \right)^{(1-\delta_k)\Lambda_k} (P_k^m)^{(1-\sigma_k)(1-\delta_k)} \Psi_k^{(1-\sigma_k)\delta_k} \times \\ &\quad \int_z \int_{\ell'} \left( z(\ell')^{\delta_k} \right)^{\Lambda_k} \widehat{\psi}_{kj}(z, \ell') d\ell' dz, \\ Avg\_Price_{Cf} &= \left( \frac{(1-\delta_C)(\sigma_C-1)}{\sigma_C} \right)^{(1-\delta_C)\Lambda_C} (P_C^m)^{(1-\sigma_C)(1-\delta_C)} \Psi_C^{(1-\sigma_C)\delta_C} \times \\ &\quad \int_z \int_{\ell'} \left( z(\ell')^{\delta_C} \right)^{\Lambda_C} (\exp(d_F \times \mathcal{I}_C^x(z, \ell')))^{-\delta_C \sigma_C \Lambda_C} \widehat{\psi}_{Cf}(z, \ell') d\ell' dz. \end{aligned}$$

So, given  $\Omega$ ,  $\Psi_C$  and  $\Psi_S$  we can compute:

$$\begin{aligned} \widetilde{Avg\_Price}_{kj} &\equiv \left( \frac{(1-\delta_k)(\sigma_k-1)}{\sigma_k} \right)^{(1-\delta_k)\Lambda_k} \Psi_k^{(1-\sigma_k)\delta_k} \int_z \int_{\ell'} \left( z(\ell')^{\delta_k} \right)^{\Lambda_k} \widehat{\psi}_{kj}(z, \ell') d\ell' dz \\ &= (P_k^m)^{(\sigma_k-1)(1-\delta_k)} Avg\_Price_{kj}, \\ \widetilde{Avg\_Price}_{Cf} &\equiv \left( \frac{(1-\delta_C)(\sigma_C-1)}{\sigma_C} \right)^{(1-\delta_C)\Lambda_C} \Psi_C^{(1-\sigma_C)\delta_C} \times \\ &\quad \int_z \int_{\ell'} \left( z(\ell')^{\delta_C} \right)^{\Lambda_C} (\exp(d_F \times \mathcal{I}_C^x(z, \ell')))^{-\delta_C \sigma_C \Lambda_C} \widehat{\psi}_{Cf}(z, \ell') d\ell' dz \\ &= (P_C^m)^{(\sigma_C-1)(1-\delta_C)} Avg\_Price_{Cf}. \end{aligned}$$

At this point, we can compute the following variables, as functions of  $M_C$  and  $M_S$

$$N_{Ci} = \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} M_C \quad (\text{S.3})$$

$$N_{Si} = \frac{\omega_{Si}}{\varrho_{Si}^{exit}} M_S \quad (\text{S.4})$$

$$N_{Cf} = \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \quad (\text{S.5})$$

$$N_{Sf} = \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S \quad (\text{S.6})$$

$$M_{Ci} = \omega_{Ci} M_C$$

$$M_{Si} = \omega_{Si} M_S$$

$$M_{Cf} = \omega_{Cf} M_C$$

$$M_{Sf} = \omega_{Sf} M_S$$

**Firm-level expenditures with sector S goods (fixed operating costs, hiring costs, entry costs, fixed export costs)**

$$\begin{aligned} E_S = & \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C (Avg\_Hiring\_Costs_{Cf} + \bar{c}_{Cf}) \\ & + \frac{\omega_{Ci} M_C}{\varrho_{Ci}^{exit}} (Avg\_Hiring\_Costs_{Ci} + \bar{c}_{Ci}) \\ & + \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S (Avg\_Hiring\_Costs_{Sf} + \bar{c}_{Sf}) \\ & + \frac{\omega_{Si} M_S}{\varrho_{Si}^{exit}} (Avg\_Hiring\_Costs_{Si} + \bar{c}_{Si}) \\ & + \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C Fraction\_Export_{Cf} f_x \\ & + M_C c_{e,C} \\ & + M_S c_{e,S} \end{aligned}$$

Define  $c_C$ :

$$\begin{aligned} c_C \equiv \frac{E_{S,C}}{M_C} = & \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} (Avg\_Hiring\_Costs_{Cf} + \bar{c}_{Cf}) \\ & + \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} (Avg\_Hiring\_Costs_{Ci} + \bar{c}_{Ci}) \\ & + \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} Fraction\_Export_{Cf} f_x \\ & + c_{e,C}, \end{aligned} \tag{S.7}$$

Where  $E_{S,C}$  is firm-level expenditures with sector  $S$  goods (fixed costs, etc) coming from  $C$ -sector activity.

Define  $c_S$ :

$$\begin{aligned} c_S \equiv \frac{E_{S,S}}{M_S} = & \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} (Avg\_Hiring\_Costs_{Sf} + \bar{c}_{Sf}) \\ & + \frac{\omega_{Si}}{\varrho_{Si}^{exit}} (Avg\_Hiring\_Costs_{Si} + \bar{c}_{Si}) \\ & + c_{e,S}, \end{aligned} \tag{S.8}$$

Where  $E_{S,S}$  is firm-level expenditures with sector  $S$  goods (fixed costs, etc) coming from  $S$ -sector activity.

We can therefore write:

$$\begin{aligned} E_S &= E_{S,C} + E_{S,S} \\ &= c_C M_C + c_S M_S \end{aligned}$$

### Market Clearing ( $C$ and $S$ sectors)

Let  $I$  denote aggregate income. Then, market clearing in the  $C$  and  $S$  sectors must lead to:

$$\begin{aligned} \zeta I + X_C^{int} &= Rev_C - Exports + \tau_a Imports \\ (1 - \zeta) I + X_S^{int} + E_S &= Rev_S \\ Imports &= Exports \end{aligned}$$

Note that expenditures on intermediates are proportional to gross revenues:

$$P_k^m \iota_k(z, \ell) = \frac{(1 - \delta_k)(\sigma_k - 1)}{\sigma_k} R_k(z, \ell),$$

which leads to:

$$\begin{aligned} X_C^{int} &= \lambda_C \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C} Rev_C \\ &\quad + \lambda_S \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S} Rev_S \\ X_S^{int} &= (1 - \lambda_C) \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C} Rev_C \\ &\quad + (1 - \lambda_S) \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S} Rev_S \end{aligned}$$

Where  $Rev_C$  and  $Rev_S$  are total gross revenues in sectors  $C$  and  $S$  respectively. Therefore:

$$\begin{aligned} I &= \left(1 - \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C}\right) Rev_C \\ &\quad + \left(1 - \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S}\right) Rev_S \\ &\quad - E_S \\ &\quad + (\tau_a - 1) Exports \end{aligned}$$

Using

$$\begin{aligned}
Rev_C &= Avg\_Revenue_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C + Avg\_Revenue_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} M_C \\
Rev_S &= Avg\_Revenue_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S + Avg\_Revenue_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} M_S \\
Exports &= Avg\_Exports_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \\
E_S &= c_C M_C + c_S M_S
\end{aligned}$$

**Step 9a:** Write aggregate income  $I$  as a function of masses of entrants  $M_C$  and  $M_S$ .

$$\begin{aligned}
I &= \left(1 - \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C}\right) \left( Avg\_Revenue_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \right. \\
&\quad \left. + Avg\_Revenue_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} M_C \right) \\
&+ \left(1 - \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S}\right) \left( Avg\_Revenue_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S \right. \\
&\quad \left. + Avg\_Revenue_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} M_S \right) \\
&- (c_C M_C + c_S M_S) \\
&+ (\tau_a - 1) \left( Avg\_Exports_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \right)
\end{aligned}$$

Therefore:

$$I = a_C M_C + a_S M_S \quad (S.9)$$

Where

$$\begin{aligned}
a_C &= \left(1 - \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C}\right) \left( Avg\_Revenue_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} \right. \\
&\quad \left. + Avg\_Revenue_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \right) \\
&+ (\tau_a - 1) \left( Avg\_Exports_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} \right) \\
&- c_C \\
a_S &= \left(1 - \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S}\right) \left( Avg\_Revenue_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} \right. \\
&\quad \left. + Avg\_Revenue_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \right) \\
&- c_S
\end{aligned}$$

**Step 9b:** Write  $P_C$  and  $P_S$  as functions of  $M_C$  and  $M_S$ .

**Price Index Sector C**

$$P_C^{1-\sigma_C} = P_{H,C}^{1-\sigma_C} + P_{F,C}^{1-\sigma_C}$$



The domestic component is given by:

$$\begin{aligned}
P_{H,C}^{1-\sigma_C} &= N_{Cf} Avg\_Price_{Cf} + N_{Ci} Avg\_Price_{Ci} \\
&= \left( \frac{\frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} Avg\_Price_{Cf} + \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} Avg\_Price_{Ci} \right) M_C \\
&= \left( \frac{\frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} \widetilde{Avg\_Price}_{Cf} (P_C^m)^{-(\sigma_C-1)(1-\delta_C)} + \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \widetilde{Avg\_Price}_{Ci} (P_C^m)^{-(\sigma_C-1)(1-\delta_C)} \right) M_C
\end{aligned}$$

We can therefore write  $P_{C,H}$  as:

$$P_{H,C}^{1-\sigma_C} = b_C^1 (P_C^m)^{(1-\sigma_C)(1-\delta_C)} M_C,$$

Where

$$b_C^1 \equiv \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} \widetilde{Avg\_Price}_{Cf} + \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \widetilde{Avg\_Price}_{Ci}$$

The foreign component is given by:

$$P_{F,C}^{1-\sigma_C} = (\epsilon \tau_a \tau_c)^{1-\sigma_C}.$$

Under Trade Balance:

$$\begin{aligned}
Exports &= \frac{D_{H,C} (\epsilon \tau_a \tau_c)^{1-\sigma_C}}{\tau_a}, \\
\Rightarrow (\epsilon \tau_a \tau_c)^{1-\sigma_C} &= \frac{\tau_a \times Exports}{D_{H,C}} \\
&= \frac{\tau_a \times N_{Cf} Avg\_Exports_{Cf}}{D_{H,C}} \\
&= \frac{\tau_a \times Avg\_Exports_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C}{\exp(\sigma_C \times d_{H,C})} \\
&= (P_C^m)^{-(\sigma_C-1)(1-\delta_C)} \frac{\tau_a \times Avg\_Exports_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C}{\Psi_C^{\frac{\sigma_C-1}{\Lambda_C}}}
\end{aligned}$$

Where we have used

$$\exp(\sigma_C \times d_{H,C}) = \left( \frac{\Psi_C}{(P_C^m)^{-(1-\delta_C)\Lambda_C}} \right)^{\frac{\sigma_C-1}{\Lambda_C}}.$$

Therefore:

$$P_{F,C}^{1-\sigma_C} = b_C^2 (P_C^m)^{(1-\sigma_C)(1-\delta_C)} M_C,$$

Where

$$b_C^2 \equiv \frac{\tau_a \times Avg\_Exports_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C}{\Psi_C^{\frac{\sigma_C-1}{\Lambda_C}}}.$$

Therefore:

$$\frac{P_C^{1-\sigma_C}}{(P_C^m)^{(1-\sigma_C)(1-\delta_C)}} = \underbrace{(b_C^1 + b_C^2)}_{b_C} M_C \quad (\text{S.10})$$

**Price Index Sector  $S$**

$$\begin{aligned} P_S^{1-\sigma_S} &= N_{Sf} \text{Avg\_Price}_{Sf} + N_{Si} \text{Avg\_Price}_{Si} \\ &= \left( \frac{\frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} \text{Avg\_Price}_{Sf} + \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \text{Avg\_Price}_{Si}} \right) M_S \\ &= \left( \frac{\frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} \widetilde{\text{Avg\_Price}}_{Sf} (P_S^m)^{-(\sigma_S-1)(1-\delta_S)} + \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \widetilde{\text{Avg\_Price}}_{Si} (P_S^m)^{-(\sigma_S-1)(1-\delta_S)}} \right) M_S \\ &\Rightarrow P_S^{1-\sigma_S} = b_S (P_S^m)^{(1-\sigma_S)(1-\delta_S)} M_S \end{aligned}$$

Where

$$b_S \equiv \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} \widetilde{\text{Avg\_Price}}_{Sf} + \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \widetilde{\text{Avg\_Price}}_{Si}$$

Therefore:

$$\frac{P_S^{1-\sigma_S}}{(P_S^m)^{(1-\sigma_S)(1-\delta_S)}} = b_S M_S. \quad (\text{S.11})$$

**Step 9c:** Write  $X_C^{int}$  as a function of  $M_C$  and  $M_S$ .

$$\begin{aligned} X_C^{int} &= \lambda_C \frac{(1-\delta_C)(\sigma_C-1)}{\sigma_C} \left( \text{Avg\_Revenue}_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} + \text{Avg\_Revenue}_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \right) M_C \\ &\quad + \lambda_S \frac{(1-\delta_S)(\sigma_S-1)}{\sigma_S} \left( \text{Avg\_Revenue}_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} + \text{Avg\_Revenue}_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \right) M_S \\ &= d_C M_C + d_S M_S \end{aligned}$$

Where

$$\begin{aligned} d_C &= \lambda_C \frac{(1-\delta_C)(\sigma_C-1)}{\sigma_C} \left( \text{Avg\_Revenue}_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} + \text{Avg\_Revenue}_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \right) \\ d_S &= \lambda_S \frac{(1-\delta_S)(\sigma_S-1)}{\sigma_S} \left( \text{Avg\_Revenue}_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} + \text{Avg\_Revenue}_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \right) \end{aligned}$$

**Step 9d:** Solve for  $\frac{M_S}{M_C}$  that matches  $\Psi_C$ .

Remember that:

$$\exp(d_{H,C}) = \left( \frac{\zeta I + X_C^{int}}{P_C^{1-\sigma_C}} \right)^{\frac{1}{\sigma_C}}$$

Using (S.9), (S.10), (S.2) and manipulating, we obtain:

$$\Psi_C^{\frac{\sigma_C-1}{\lambda_C}} = \left( \frac{\zeta a_C}{b_C} + \frac{d_C}{b_C} \right) + \frac{\zeta a_S + d_S}{b_C} \frac{M_S}{M_C}$$

$$\frac{M_S}{M_C} = \frac{b_C}{\zeta a_S + d_S} \left( \Psi_C^{\frac{\sigma_C-1}{\lambda_C}} - \left( \frac{\zeta a_C}{b_C} + \frac{d_C}{b_C} \right) \right)$$

**Step 9e:** Separately pin down  $M_C$  and  $M_S$  using the labor market clearing equation  $\bar{L} - L_u = \sum_{k=C,S,j=i,f} L_{kj}$ . Express  $M_C$  and  $M_S$  as functions of  $L_u$ .

To separately pin down  $M_C$  and  $M_S$ , use the labor market clearing equation.

$$\begin{aligned} \bar{L} - L_u &= N_{Cf} \text{Avg\_Size}_{Cf} + N_{Ci} \text{Avg\_Size}_{Ci} + N_{Sf} \text{Avg\_Size}_{Sf} + N_{Si} \text{Avg\_Size}_{Si} \\ &= \frac{\varrho_{Ci}^{\text{change}} \omega_{Ci} + \varrho_{Ci}^{\text{exit}} \omega_{Cf}}{\varrho_{Cf}^{\text{exit}} \varrho_{Ci}^{\text{exit}}} M_C \text{Avg\_Size}_{Cf} + \frac{\omega_{Ci} M_C}{\varrho_{Ci}^{\text{exit}}} \text{Avg\_Size}_{Ci} + \\ &\quad \frac{\varrho_{Si}^{\text{change}} \omega_{Si} + \varrho_{Si}^{\text{exit}} \omega_{Sf}}{\varrho_{Sf}^{\text{exit}} \varrho_{Si}^{\text{exit}}} M_S \text{Avg\_Size}_{Sf} + \frac{\omega_{Si} M_S}{\varrho_{Si}^{\text{exit}}} \text{Avg\_Size}_{Si} \\ &= \left( \frac{\varrho_{Ci}^{\text{change}} \omega_{Ci} + \varrho_{Ci}^{\text{exit}} \omega_{Cf}}{\varrho_{Cf}^{\text{exit}} \varrho_{Ci}^{\text{exit}}} \text{Avg\_Size}_{Cf} + \frac{\omega_{Ci}}{\varrho_{Ci}^{\text{exit}}} \text{Avg\_Size}_{Ci} \right) M_C + \\ &\quad \left( \frac{\varrho_{Si}^{\text{change}} \omega_{Si} + \varrho_{Si}^{\text{exit}} \omega_{Sf}}{\varrho_{Sf}^{\text{exit}} \varrho_{Si}^{\text{exit}}} \text{Avg\_Size}_{Sf} + \frac{\omega_{Si}}{\varrho_{Si}^{\text{exit}}} \text{Avg\_Size}_{Si} \right) M_S \end{aligned}$$

At this point, we can only express  $M_C$  and  $M_S$  as functions of  $L_u$ .

From now on write

$$\begin{aligned} \left( \frac{M_S}{M_C} \right)^* &= \frac{b_C}{\zeta a_S + d_S} \left( \Psi_C^{\frac{\sigma_C-1}{\lambda_C}} - \left( \frac{\zeta a_C}{b_C} + \frac{d_C}{b_C} \right) \right) \\ \Rightarrow M_S &= \underbrace{\frac{b_C}{\zeta a_S + d_S} \left( \Psi_C^{\frac{\sigma_C-1}{\lambda_C}} - \left( \frac{\zeta a_C}{b_C} + \frac{d_C}{b_C} \right) \right)}_{AA} M_C \end{aligned}$$

Therefore:

$$\begin{aligned} M_S &= AA \times M_C \\ AA &= \frac{b_C}{\zeta a_S + d_S} \left( \Psi_C^{\frac{\sigma_C-1}{\lambda_C}} - \left( \frac{\zeta a_C}{b_C} + \frac{d_C}{b_C} \right) \right) \end{aligned} \tag{S.12}$$

So that:

$$\begin{aligned} \bar{L} - L_u &= \left( \frac{\varrho_{Ci}^{\text{change}} \omega_{Ci} + \varrho_{Ci}^{\text{exit}} \omega_{Cf}}{\varrho_{Cf}^{\text{exit}} \varrho_{Ci}^{\text{exit}}} \text{Avg\_Size}_{Cf} + \frac{\omega_{Ci}}{\varrho_{Ci}^{\text{exit}}} \text{Avg\_Size}_{Ci} \right) M_C + \\ &\quad \left( \frac{\varrho_{Si}^{\text{change}} \omega_{Si} + \varrho_{Si}^{\text{exit}} \omega_{Sf}}{\varrho_{Sf}^{\text{exit}} \varrho_{Si}^{\text{exit}}} \text{Avg\_Size}_{Sf} + \frac{\omega_{Si}}{\varrho_{Si}^{\text{exit}}} \text{Avg\_Size}_{Si} \right) AA \times M_C \\ &= BB \times M_C \end{aligned}$$

$$BB = \left( \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} Avg\_Size_{Cf} + \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} Avg\_Size_{Ci} \right) + \left( \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} Avg\_Size_{Sf} + \frac{\omega_{Si}}{\varrho_{Si}^{exit}} Avg\_Size_{Si} \right) AA \quad (S.13)$$

Finally:

$$M_C = \frac{\bar{L} - L_u}{BB} \quad (S.14)$$

$$M_S = \frac{AA}{BB} (\bar{L} - L_u) \quad (S.15)$$

**Step 9f:** Express masses of firms  $N_{kj}$  as functions of  $L_u$ .

Substituting (S.14) and (S.15) into (S.3)-(S.6) to obtain the masses of firms:

$$\begin{aligned} N_{Ci} &= \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} M_C = \underbrace{\frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \frac{1}{BB}}_{EE_C} (\bar{L} - L_u) = EE_C (\bar{L} - L_u) \\ N_{Si} &= \frac{\omega_{Si}}{\varrho_{Si}^{exit}} M_S = \underbrace{\frac{\omega_{Si}}{\varrho_{Si}^{exit}} \frac{AA}{BB}}_{EE_S} (\bar{L} - L_u) = EE_S (\bar{L} - L_u) \\ N_{Cf} &= \underbrace{\frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} \frac{1}{BB}}_{DD_C} (\bar{L} - L_u) \\ N_{Sf} &= \underbrace{\frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} \frac{AA}{BB}}_{DD_S} (\bar{L} - L_u) \end{aligned}$$

**Step 9g:** Express aggregate posted vacancies  $V_{kj}$  as functions of  $L_u$ .

Now, substituting the expressions for the  $N_{kj}$ 's to obtain the number of vacancies in each sector as a function of  $L_u$ :

$$\begin{aligned} V_{Cf} &= N_{Cf} Avg\_Vacancies_{Cf} + \frac{\omega_{Cf} M_C}{\mu^v} \quad (S.16) \\ &= Avg\_Vacancies_{Cf} \times DD_C (\bar{L} - L_u) + \frac{\omega_{Cf}}{\mu^v} \frac{1}{BB} (\bar{L} - L_u) \\ &= \underbrace{\left( Avg\_Vacancies_{Cf} \times DD_C + \frac{\omega_{Cf}}{\mu^v} \frac{1}{BB} \right)}_{FF_C} (\bar{L} - L_u) \\ &= FF_C \times (\bar{L} - L_u) \end{aligned}$$

$$\begin{aligned}
V_{Ci} &= N_{Ci} \text{Avg\_Vacancies}_{Ci} + \frac{\omega_{Ci} M_C}{\mu^v} \\
&= \text{Avg\_Vacancies}_{Ci} \times EE_C (\bar{L} - L_u) + \frac{\omega_{Ci}}{\mu^v} \frac{1}{BB} (\bar{L} - L_u) \\
&= \underbrace{\left( \text{Avg\_Vacancies}_{Ci} \times EE_C + \frac{\omega_{Ci}}{\mu^v} \frac{1}{BB} \right)}_{GG_C} (\bar{L} - L_u) \\
&= GG_C \times (\bar{L} - L_u)
\end{aligned} \tag{S.17}$$

$$\begin{aligned}
V_{Sf} &= N_{Sf} \text{Avg\_Vacancies}_{Sf} + \frac{\omega_{Sf} M_S}{\mu^v} \\
&= \underbrace{\left( \text{Avg\_Vacancies}_{Sf} \times DD_S + \frac{\omega_{Sf}}{\mu^v} \frac{AA}{BB} \right)}_{FF_S} (\bar{L} - L_u) \\
&= FF_S \times (\bar{L} - L_u)
\end{aligned} \tag{S.18}$$

$$\begin{aligned}
V_{Si} &= N_{Si} \text{Avg\_Vacancies}_{Si} + \frac{\omega_{Si} M_S}{\mu^v} \\
&= \underbrace{\left( \text{Avg\_Vacancies}_{Si} \times EE_S + \frac{\omega_{Si}}{\mu^v} \frac{AA}{BB} \right)}_{GG_S} (\bar{L} - L_u) \\
&= GG_S \times (\bar{L} - L_u)
\end{aligned} \tag{S.19}$$

$$\begin{aligned}
\tilde{V} &= V_{Cf} + V_{Ci} + V_{Sf} + V_{Si} \\
&= \underbrace{(FF_C + GG_C + FF_S + GG_S)}_{JJ} \times (\bar{L} - L_u) \\
&= JJ \times (\bar{L} - L_u)
\end{aligned}$$

**Step 9h:** Use equation for  $\mu^v$  to obtain  $L_u$ .

We have written each  $V_{kj}$  in terms of  $L_u$ . Now, note that

$$\mu^v = \phi \left( \frac{L_u}{\tilde{V}} \right)^{1-\xi}$$

We can invert this equation to obtain  $L_u$ .

$$\begin{aligned}
\mu^v &= \phi \left( \frac{L_u}{JJ \times (\bar{L} - L_u)} \right)^{1-\xi} \\
\Rightarrow L_u^* &= \frac{(\mu^v)^{\frac{1}{1-\xi}} \times JJ \times \bar{L}}{\phi^{\frac{1}{1-\xi}} + (\mu^v)^{\frac{1}{1-\xi}} \times JJ}
\end{aligned}$$

**Step 9i:** Go back and obtain masses of entrants  $M_k$ 's (equations (S.14) and (S.15)), masses of firms  $N_{kj}$ 's (equations (S.3)-(S.6)), and aggregate vacancies  $V_{kj}$ 's (equations (S.16)-(S.19)). We are now able to compute transitions out of unemployment  $\mu_{kj}^e$  (Step 8).

**Step 9j:** Recover price indices  $P_C$  and  $P_S$ .

Equations (S.1) and (S.10) lead to:

$$P_C = \left( (b_C M_C)^{\frac{1}{1-\sigma_C}} \left( \frac{1}{\lambda_C^{\lambda_C} (1-\lambda_C)^{1-\lambda_C}} \right)^{(1-\delta_C)} \right)^{\frac{1}{1-(1-\delta_C)\lambda_C}} P_S^{\frac{(1-\delta_C)(1-\lambda_C)}{1-(1-\delta_C)\lambda_C}}$$

Defining

$$\varpi_C = \left( (b_C M_C)^{\frac{1}{1-\sigma_C}} \left( \frac{1}{\lambda_C^{\lambda_C} (1-\lambda_C)^{1-\lambda_C}} \right)^{(1-\delta_C)} \right)^{\frac{1}{1-(1-\delta_C)\lambda_C}}$$

and

$$\varkappa_C = \frac{(1-\delta_C)(1-\lambda_C)}{1-(1-\delta_C)\lambda_C}$$

Allows us to write

$$P_C = \varpi_C P_S^{\varkappa_C}$$

Equations (S.1) and (S.11) lead to:

$$P_S = \left( (b_S M_S)^{\frac{1}{1-\sigma_S}} \left( \frac{1}{\lambda_S^{\lambda_S} (1-\lambda_S)^{1-\lambda_S}} \right)^{(1-\delta_S)} \right)^{\frac{1}{1-(1-\delta_S)(1-\lambda_S)}} P_C^{\frac{(1-\delta_S)\lambda_S}{1-(1-\delta_S)(1-\lambda_S)}}$$

Writing

$$\varpi_S = \left( (b_S M_S)^{\frac{1}{1-\sigma_S}} \left( \frac{1}{\lambda_S^{\lambda_S} (1-\lambda_S)^{1-\lambda_S}} \right)^{(1-\delta_S)} \right)^{\frac{1}{1-(1-\delta_S)(1-\lambda_S)}}$$

and

$$\varkappa_S = \frac{(1-\delta_S)\lambda_S}{1-(1-\delta_S)(1-\lambda_S)}$$

Allows us to write

$$P_S = \varpi_S P_C^{\varkappa_S}$$

Solving the system leads to:

$$P_C = (\varpi_C (\varpi_S)^{\varkappa_C})^{\frac{1}{1-\varkappa_S \varkappa_C}}$$

**Step 9k:** Compute deviation between government revenues and spending with unemployment insurance  $Dev_T$ .

## Government Revenue

$$\begin{aligned}
G_{Rev} = & \frac{\sigma_C - (1 - \delta_C)(\sigma_C - 1)}{\sigma_C} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \tau_y Avg\_Revenue_{Cf} \\
& + \frac{\sigma_S - (1 - \delta_S)(\sigma_S - 1)}{\sigma_S} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S \tau_y Avg\_Revenue_{Sf} \\
& + \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \tau_w Avg\_wbill_{Cf} \\
& + \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S \tau_w Avg\_wbill_{Sf} \\
& + \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C Avg\_Firing\_Costs_{Cf} \\
& + \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S Avg\_Firing\_Costs_{Sf} \\
& + (\tau_a - 1) \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C Avg\_Exports_{Cf}
\end{aligned}$$

## Government Spending with Unemployment Insurance

$$G_{UI} = \underbrace{\left( b^u \times \underbrace{\sum_k (W_{kf}^{DS} + W_{kf}^{EE} + W_{kf}^D)}_{\text{mass of formal workers who transition to unemployment}} \right)}_{\text{Total Expenditure with Unemployment Benefits}}$$

## Government Transfers

$$T = G_{Rev} - G_{UI}$$

We impose in the objective function that  $Dev_T \geq 0$ —in other words, we highly penalize  $Dev_T < 0$

$$Dev_T = \frac{G_{Rev} - G_{UI}}{G_{Rev}}$$

When we compute aggregate income, we implicitly assumed that  $G_{Rev} - G_{UI} \geq 0$ .

# II Simulation

## II.1 Simulation Algorithm

Fix  $P_S$  at  $\bar{P}_S$ . Write the value added function as:

$$VA_k(z, \ell) = \Theta_k \left( \frac{\bar{P}_S^{1-\lambda_k}}{\lambda_k^{\lambda_k} (1-\lambda_k)^{1-\lambda_k}} \right)^{-(1-\delta_k)\Lambda_k} P_C^{-\lambda_k(1-\delta_k)\Lambda_k} (\exp(d_{H,k}))^{\frac{\sigma_k}{\sigma_k-1}\Lambda_k} (z\ell^{\delta_k})^{\Lambda_k}$$

Define

$$\Xi_k \equiv \Theta_k \left( \frac{\bar{P}_S^{1-\lambda_k}}{\lambda_k^{\lambda_k} (1-\lambda_k)^{1-\lambda_k}} \right)^{-(1-\delta_k)\Lambda_k},$$

and

$$\Phi_k \equiv P_C^{-\lambda_k(1-\delta_k)\Lambda_k} (\exp(d_{H,k}))^{\frac{\sigma_k}{\sigma_k-1}\Lambda_k}.$$

Rewrite the value added function as:

$$VA_k(z, \ell) = \Xi_k \Phi_k (z \ell^{\delta_k})^{\Lambda_k}.$$

$\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S$  are the endogenous variables to be determined in equilibrium. For a given value of these variables, Steps 1 through 11 below compute the deviations from equilibrium conditions given by  $L_i(\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S)$  for  $i = 1, \dots, 5$ . We then need to find values  $(\mu^v)^*, \vartheta_{J_u}^*, d_F^*, \Phi_C^*, \Phi_S^*$  solving  $L_i((\mu^v)^*, \vartheta_{J_u}^*, d_F^*, \Phi_C^*, \Phi_S^*) = 0$  for all  $i = 1, \dots, 5$ . We discuss potential solutions to this problem in Step 12.

We proceed by first imposing values for  $\vartheta_{J_u}, \mu^v, d_F, \Phi_C, \Phi_S$ .

**Step 1:** This step solves for wage schedules  $w_{kf}(z, \ell')$ ,  $w_{ki}(z, \ell')$  as well as value functions  $V_{kf}(z, \ell)$ ,  $V_{ki}(z, \ell)$ ,  $J_{kf}^e(z, \ell')$ ,  $J_{ki}^e(z, \ell')$ , and firms' policy functions.

**Step 1a:** Compute value added functions  $VA_k(z, \ell)$ .

**Step 1b:** Compute wage schedules  $w_{kf}(z, \ell')$

- Guess a wage schedule  $w_{kf}(z, \ell')$
- Compute the resulting  $V_{kf}(z, \ell')$  using (14)
- Compute  $J_{kf}^e(z, \ell')$  using (I.1)
- Compute  $w_{kf}^u(z, \ell')$  using equation (28)
- Let  $\hat{w}_{kf}^u(z, \ell') = \omega_0 + \omega_1 \frac{VA_k(z, \ell')}{\ell'}$  be the linear projection of  $w_{kf}^u(z, \ell')$  on  $\left[1, \frac{VA_k(z, \ell')}{\ell'}\right]$
- Update  $w_{kf}(z, \ell') = \max \left\{ \hat{w}_{kf}^u(z, \ell'), b_u + \vartheta_{J_u} - \frac{1}{1+r} J_{kf}^e(z, \ell'), \underline{w} \right\}$
- Restart until convergence

**Step 1c:** Compute wage schedules  $w_{ki}(z, \ell')$

- Guess a wage schedule  $w_{ki}(z, \ell')$
- Compute the resulting  $V_{ki}(z, \ell')$  using (18)
- Compute  $J_{ki}^e(z, \ell')$  using (I.2)
- Compute  $w_{ki}^u(z, \ell')$  using equation (31)
- Let  $\hat{w}_{ki}^u(z, \ell') = \omega_0 + \omega_1 \left(1 - \frac{\sigma_k}{\sigma_k - (1-\delta_k)(\sigma_k-1)} p_{ki}(\ell')\right) \frac{VA_k(z, \ell')}{\ell'}$  be the linear projection of  $w_{ki}^u(z, \ell')$  on  $\left[1, \left(1 - \frac{\sigma_k}{\sigma_k - (1-\delta_k)(\sigma_k-1)} p_{ki}(\ell')\right) \frac{VA_k(z, \ell')}{\ell'}\right]$



- Update  $w_{ki}(z, \ell') = \max \left\{ \hat{w}_{ki}^u(z, \ell'), \vartheta_{J_u} - \frac{1}{1+r} J_{ki}^e(z, \ell') \right\}$
- Restart until convergence

**Step 2:** Solve for firms' entry decisions. Compute the fraction of entrants in the formal and informal sectors as follows:

$$\omega_{kf} \equiv \Pr \left( I_k^{formal}(z) = 1 \right) = \int_z I_k^{formal}(z) g_k^e(z) dz$$

$$\omega_{ki} \equiv \Pr \left( I_k^{informal}(z) = 1 \right) = \int_z I_k^{informal}(z) g_k^e(z) dz$$

Therefore, if  $M_k$  is the mass of entrants in sector  $k$ , the masses of formal and informal entrants in sector  $k$  are given by:

$$M_{ki} = \omega_{ki} M_k$$

$$M_{kf} = \omega_{kf} M_k$$

Finally, compute the distribution of  $z$  productivities among entrants, conditional on entry into sector  $kj$ .

$$\psi_{ki}^e(z) = \frac{g_k^e(z) I_k^{informal}(z)}{\int_{\tilde{z}} g_k^e(\tilde{z}) I_k^{informal}(\tilde{z}) d\tilde{z}},$$

$$\psi_{kf}^e(z) = \frac{g_k^e(z) I_k^{formal}(z)}{\int_{\tilde{z}} g_k^e(\tilde{z}) I_k^{formal}(\tilde{z}) d\tilde{z}}.$$

**Step 3:** Compute the steady-state distribution of states. For informal firms, start with a guess for  $\psi_{ki}$ . Then, compute

$$\varrho_{ki}^{exit} = \alpha_k + (1 - \alpha_k) \int_z \int_{\ell} \left( I_{ki}^{exit}(z, \ell) + I_{ki}^{change}(z, \ell) \right) \psi_{ki}(z, \ell) d\ell dz.$$

In steady state  $N_{ki} = (1 - \varrho_{ki}^{exit}) N_{ki} + M_{ki}$ . Therefore, set  $\frac{M_{ki}}{N_{ki}}$ , the fraction of sector  $k$  informal firms that are entrants, to:

$$\boxed{\frac{M_{ki}}{N_{ki}} = \varrho_{ki}^{exit} = \frac{\omega_{ki} M_k}{N_{ki}}}.$$

Now, compute  $\tilde{\psi}_{ki}$ :

$$\tilde{\psi}_{ki}(z, \ell) = \mathcal{I}[\ell = 1] \times \varrho_{ki}^{exit} \times \psi_{ki}^e(z) \\ + \mathcal{I}[\ell \geq 1] \times (1 - \alpha_k) \times \psi_{ki}(z, \ell) I_{ki}^{stay}(z, \ell),$$

and  $\hat{\psi}_{ki}$ :

$$\hat{\psi}_{ki}(z, \ell') = \int_{\ell} \tilde{\psi}_{ki}(z, \ell) \mathcal{I}(L_{ki}(z, \ell) = \ell') d\ell$$

Update  $\psi_{ki}$  with:

$$\psi_{ki}(z', \ell') = \int_z \hat{\psi}_{ki}(z, \ell') g_k(z'|z) dz,$$

and repeat until convergence of  $\psi_{ki}$ . This converged value of  $\psi_{ki}$  will be used directly in the computation of  $\psi_{kf}$  below.

For formal firms, start with guess for  $\psi_{kf}$  and compute:

$$\varrho_{kf}^{exit} = \alpha_k + (1 - \alpha_k) \int_z \int_\ell I_{kf}^{exit}(z, \ell) \psi_{kf}(z, \ell) d\ell dz,$$

$$\varrho_{ki}^{change} = (1 - \alpha_k) \int_z \int_\ell I_{ki}^{change}(z, \ell) \psi_{ki}(z, \ell) d\ell dz.$$

In steady state:

$$\begin{aligned} \varrho_{kf}^{exit} N_{kf} &= \varrho_{ki}^{change} \underbrace{N_{ki}}_{\frac{\omega_{ki} M_k}{\varrho_{ki}^{exit}}} + \omega_{kf} M_k \\ &= M_k \left( \frac{\varrho_{ki}^{change}}{\varrho_{ki}^{exit}} \omega_{ki} + \omega_{kf} \right) \end{aligned}$$

So that:

$$\boxed{\frac{M_{kf}}{N_{kf}} = \frac{M_k \omega_{kf}}{N_{kf}} = \frac{\varrho_{kf}^{exit} \omega_{kf}}{\frac{\varrho_{ki}^{change}}{\varrho_{ki}^{exit}} \omega_{ki} + \omega_{kf}}}$$

Also, note that

$$\frac{M_{kf}}{N_{kf}} \times \frac{N_{ki}}{M_{ki}} = \frac{\varrho_{kf}^{exit} \omega_{kf}}{\frac{\varrho_{ki}^{change}}{\varrho_{ki}^{exit}} \omega_{ki} + \omega_{kf}} \frac{1}{\varrho_{ki}^{exit}} = \frac{\varrho_{kf}^{exit} \omega_{kf}}{\varrho_{ki}^{change} \omega_{ki} + \varrho_{ki}^{exit} \omega_{kf}}$$

and

$$\frac{M_{kf}}{N_{kf}} \times \frac{N_{ki}}{M_{ki}} = \frac{\omega_{kf}}{\omega_{ki}} \frac{N_{ki}}{N_{kf}}$$

Therefore,

$$\boxed{\frac{N_{ki}}{N_{kf}} = \frac{\varrho_{kf}^{exit} \omega_{ki}}{\varrho_{ki}^{change} \omega_{ki} + \varrho_{ki}^{exit} \omega_{kf}}}$$

Compute  $\tilde{\psi}_{kf}$  as:

$$\begin{aligned} \tilde{\psi}_{kf}(z, \ell) &= \mathcal{I}[\ell = 1] \times \underbrace{\frac{\varrho_{kf}^{exit} \omega_{kf}}{\frac{\varrho_{ki}^{change}}{\varrho_{ki}^{exit}} \omega_{ki} + \omega_{kf}}}_{\frac{M_{kf}}{N_{kf}}} \psi_{kf}^e(z) \\ &+ \mathcal{I}[\ell \geq 1] \times \left( (1 - \alpha_k) \psi_{kf}(z, \ell) I_{kf}^{stay}(z, \ell) \right. \\ &\quad \left. + (1 - \alpha_k) \underbrace{\frac{\varrho_{kf}^{exit} \omega_{ki}}{\varrho_{ki}^{change} \omega_{ki} + \varrho_{ki}^{exit} \omega_{kf}}}_{\frac{N_{ki}}{N_{kf}}} \psi_{ki}(z, \ell) I_{ki}^{change}(z, \ell) \right) \end{aligned}$$

and  $\hat{\psi}_{kf}$  as:

$$\hat{\psi}_{kf}(z, \ell') = \int_\ell \tilde{\psi}_{kf}(z, \ell) \mathcal{I}(L_{kf}(z, \ell) = \ell') d\ell.$$

Update  $\psi_{kf}$  with:

$$\psi_{kf}(z', \ell') = \int_z \hat{\psi}_{kf}(z, \ell') g_k(z'|z) dz,$$

and repeat until convergence of  $\psi_{kf}$ .

At this point we have the following objects:  $\psi_{kj}$ ,  $\tilde{\psi}_{kj}$ ,  $\hat{\psi}_{kj}$ ,  $\varrho_{ki}^{exit}$ ,  $\varrho_{ki}^{change}$ ,  $\varrho_{kf}^{exit}$ ,  $\varrho_{kf}^{change}$ ,  $\chi_{ki \rightarrow f}^{change}$ ,  $\chi_{kf}^{layoff}$ , and  $\chi_{ki}^{leave}$  (see equations (C.7), (C.11) and (C.13)).

**Step 4:** Compute the values of entry  $V_k^e$  ( $k = C, S$ ):

$$V_k^e = \int_z \left[ V_{ki}^e(z) I_k^{informal}(z) + V_{kf}^e(z) I_k^{formal}(z) \right] g_k^e(z) dz$$

and compute the deviations

$$L_5(\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S) = L_5(\vartheta_{J_u}, \mu^v, \Phi_S) = Dev_{entry, S} = \frac{V_S^e - c_{e, S}}{c_{e, S}}$$

$$L_4(\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S) = L_4(\vartheta_{J_u}, \mu^v, d_F, \Phi_C) = Dev_{entry, C} = \frac{V_C^e - c_{e, C}}{c_{e, C}}$$

**Step 5:** This step solves for masses of entrants  $M_k$ 's, masses of firms  $N_{kj}$ 's, aggregate vacancies  $V_{kj}$ 's consistent with  $\Phi_C$ ,  $\Phi_S$ ,  $\vartheta_{J_u}$ ,  $d_F$  and  $\mu^v$ .

**Step 5a:** Write aggregate income  $I$  as a function of masses of entrants  $M_C$  and  $M_S$ .

**Step 5b:** Write  $M_C$  as a function of  $P_C$  and  $M_S$  as a function of  $M_C$  and  $\bar{P}_S$ .

**Step 5c:** Write  $X_C^{int}$  as a function of  $M_C$  and  $M_S$

**Step 5d:** Pin down  $M_C$  using the equation defining  $\Phi_C$ , then obtain  $M_S$ .

**Step 5e:** Obtain masses of firms  $N_{kj}$ .

**Step 5f:** Obtain aggregate posted vacancies  $V_{kj}$  and  $\tilde{V}$ .

**Step 5g:** Save the values for  $P_C$  and  $P_{F, C}$  to be used in Step 9.

**Step 6:** Compute  $L_u$

$$L_u = \left( \frac{\mu^v}{\phi} \right)^{\frac{1}{1-\xi}} \tilde{V}$$

**Step 7:** Obtain job finding rates  $\mu_{kj}^e$  using aggregate vacancies  $V_{kj}$ 's and mass of unemployment  $L_u$  obtained in Steps 5 and 6.

$$\mu_{kj}^e = \frac{m_{kj}}{L_u} = \phi \frac{V_{kj}}{\tilde{V}} \left( \frac{\tilde{V}}{L_u} \right)^\xi$$

**Step 8:** Use equations (C.14)-(C.15) to obtain allocations  $L_{Cf}$ ,  $L_{Ci}$ ,  $L_{Sf}$ ,  $L_{Si}$ .

$$\begin{aligned} L_{Ci} &= \frac{\mu_{Ci}^e L_u}{\chi_{Ci}^{leave}} \\ L_{Si} &= \frac{\mu_{Si}^e L_u}{\chi_{Si}^{leave}} \\ L_{Cf} &= \frac{\mu_{Cf}^e L_u + \chi_{Ci \rightarrow f}^{change} L_{Ci}}{\chi_{Cf}^{layoff}} \\ L_{Sf} &= \frac{\mu_{Sf}^e L_u + \chi_{Si \rightarrow f}^{change} L_{Si}}{\chi_{Sf}^{layoff}} \end{aligned}$$

**Step 9:** Compute

$$\bar{\epsilon} = \frac{P_{F,C}}{\tau_a \tau_c},$$

where  $P_{F,C}$  was determined in Step 5.

Compute:

$$d'_F = \log \left( \left( 1 + \frac{D_F^*}{\exp(\sigma_C \times d_{H,C})} \bar{\epsilon}^{\sigma_C} \tau_c^{1-\sigma_C} \right)^{\frac{1}{\sigma_C}} \right),$$

where

$$\exp(\sigma_C \times d_{H,C}) = \Phi_C^{\frac{\sigma_C-1}{\lambda_C}} (P_C)^{\lambda_C(1-\delta_C)(\sigma_C-1)},$$

and  $P_C$  was determined in Step 5. Compute the deviation

$$L_3(\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S) = Dev_{d_F} = \frac{d_F - d'_F}{d_F}$$

**Step 10:** Compute deviation from the labor market clearing equation:

$$L_1(\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S) = Dev_L = \frac{\bar{L} - (L_{Cf} + L_{Ci} + L_{Sf} + L_{Si} + L_u)}{\bar{L}}$$

**Step 11:** Compute the deviation

$$\begin{aligned} L_2(\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S) &= Dev_{J_u} \\ &= 1 - \frac{\left( \sum_{k,j} \mu_{kj}^e \int_{\ell} \int_z \bar{J}_{kj}^e(z, L_{kj}(z, \ell)) g_{kj}(z, \ell) dz d\ell + \left( 1 - \sum_{k,j} \mu_{kj}^e \right) \vartheta_{J_u} \right)}{(1+r)(\vartheta_{J_u} - b)} \end{aligned}$$

Therefore, given  $\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S$ , we can compute deviations  $L_1, L_2, L_3, L_4, L_5$ .

**Step 12:** The equilibrium is given by  $(\mu^v)^*, \vartheta_{J_u}^*, d_F^*, \Phi_C^*, \Phi_S^*$  solving

$$L_i((\mu^v)^*, \vartheta_{J_u}^*, d_F^*, \Phi_C^*, \Phi_S^*) = 0 \text{ for all } i = 1, \dots, 5$$

**Step 13:** Compute the price index for exports

$$P_X^* \equiv \left( \int_{N_{F,C}}^{N_C} \mathcal{I}_C^x(n) p_x^*(n)^{1-\sigma_C} dn \right)^{\frac{1}{1-\sigma_C}}$$

Note that

$$Exports = \epsilon D_F^* (P_X^*)^{1-\sigma_C}$$

So that:

$$P_X^* = \left( \frac{Exports}{\epsilon D_F^*} \right)^{\frac{1}{1-\sigma_C}}$$

A key difficulty is that, given the discrete approximations for the state space, the system above has discontinuities. We list a few solutions we implemented.

- Solve for the system using a sequential bisection method. This procedure has the drawback of being very slow.
- Solve for the system using an optimization routine minimizing the norm of the system. This procedure has the drawback of also being slow and to potentially be stuck in local minima.
- Our preferred solution is to approximate each function  $L_i(\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S)$  with a third degree polynomial on the arguments. To do so, we draw a large number of values for  $(\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S)$  and follow Steps 1 through 11 above to compute  $L_i(\mu^v, \vartheta_{J_u}, d_F, \Phi_C, \Phi_S)$  at each of these points. We then fit third degree polynomials for each  $L_i$  function  $i = 1, \dots, 5$ . Finally, we can use an out-of-the shelf solver to find the root of this approximated system.

## II.2 Simulation Algorithm – Details

This section details the steps within Step 5 of the estimation procedure.

**Step 5:** This step solves for masses of entrants  $M_k$ 's, masses of firms  $N_{kj}$ 's, aggregate vacancies  $V_{kj}$ 's consistent with  $\Phi_C, \Phi_S, \vartheta_{J_u}, d_F$ , and  $\mu^v$ .

We start with some definitions... Averages "per firm". All these quantities can be computed after Step 4, that is, after solving for the steady state distribution of states.

$$Avg\_wbill_{ki} = \int_z \int_{\ell'} [w_{ki}(z, \ell') \ell'] \hat{\psi}_{ki}(z, \ell') d\ell' dz \text{ for } k = C, S$$

$$Avg\_wbill_{kf} = \int_z \int_{\ell'} [\max\{w_{kf}(z, \ell'), \underline{w}\} \ell'] \hat{\psi}_{kf}(z, \ell') d\ell' dz \text{ for } k = C, S$$

$$Avg\_Firing\_Costs_{kf} = \kappa \int_z \int_{\ell} [(\ell - L_{kf}(z, \ell)) (1 - I_{kf}^{hire}(z, \ell))] \tilde{\psi}_{kf}(z, \ell) d\ell dz \text{ for } k = C, S$$

$$Avg\_Hiring\_Costs_{kj} = \int_z \int_{\ell} [H_{kj}(\ell, L_{kj}(z, \ell)) I_{kj}^{hire}(z, \ell)] \tilde{\psi}_{kj}(z, \ell) d\ell dz \text{ for } k = C, S; j = i, f$$

$$Avg\_Revenue_{kj} = \int_z \int_{\ell'} R_k(z, \ell') \hat{\psi}_{kj}(z, \ell') d\ell' dz \text{ for } k = C, S; j = i, f$$

$$Avg\_InfPenalty_{ki} = \int_z \int_{\ell'} [p_{ki}(\ell') R_k(z, \ell')] \hat{\psi}_{ki}(z, \ell') d\ell' dz \text{ for } k = C, S$$

$$Avg\_Vacancies_{kj} = \int_z \int_{\ell} v_{kj}(z, \ell) \tilde{\psi}_{kj}(z, \ell) d\ell dz \text{ for } k = C, S; j = i, f$$

$$Avg\_Exports_{Cf} = (1 - \exp(-\sigma_C \times d_F)) \int_z \int_{\ell'} [R_C(z, \ell') \mathcal{I}_C^x(z, \ell')] \hat{\psi}_{Cf}(z, \ell') d\ell' dz$$

$$Fraction\_Export_{Cf} = \int_z \int_{\ell'} \mathcal{I}_C^x(z, \ell') \hat{\psi}_{Cf}(z, \ell') d\ell' dz$$

$$Avg\_size_{kj} = \int_z \int_{\ell} \ell \psi_{kj}(z, \ell) d\ell dz \text{ for } k = C, S; j = i, f$$

Now, define

$$\begin{aligned} Avg\_Price_{kj} &= \int_z \int_{\ell'} p_{kj}(z, \ell')^{1-\sigma_k} \widehat{\psi}_{kj}(z, \ell') d\ell' dz \\ &= \int_z \int_{\ell'} \left( \frac{R_k(z, \ell')}{q_k(z, \ell', \iota_k(z, \ell'))} \right)^{1-\sigma_k} \widehat{\psi}_{kj}(z, \ell') d\ell' dz \text{ for } k = C, S; j = i, f. \end{aligned}$$

We cannot compute  $Avg\_Price_{kj}$ —given  $\Omega$ ,  $\Phi_C$  and  $\Phi_S$ . However, note that:

$$\begin{aligned} Avg\_Price_{kj} &= \widetilde{\Xi}_k P_C^{(1-\sigma_k)(1-\delta_k)\lambda_k} \Phi_k^{\delta_k(1-\sigma_k)} \int \int \left( z(\ell')^{\delta_k} \right)^{\Lambda_k} \widehat{\psi}_{kj}(z, \ell') dz d\ell' \\ Avg\_Price_{Cf} &= \widetilde{\Xi}_C P_C^{(1-\sigma_C)(1-\delta_C)\lambda_C} \Phi_C^{\delta_C(1-\sigma_C)} \int \int \left( z(\ell')^{\delta_C} \right)^{\Lambda_C} (\exp(d_F \times \mathcal{I}_C^x(z, \ell')))^{-\delta_C \sigma_C \Lambda_C} \widehat{\psi}_{Cf}(z, \ell') dz d\ell' \\ \widetilde{\Xi}_k &= \left( \frac{\overline{P}_S^{1-\lambda_k}}{\lambda_k^{\lambda_k} (1-\lambda_k)^{1-\lambda_k}} \right)^{-(1-\delta_k)\Lambda_k} \left( \frac{(1-\delta_k)(\sigma_k-1)}{\sigma_k} \right)^{(1-\delta_k)\Lambda_k} \end{aligned}$$

So, given  $\Omega$ ,  $\Phi_C$  and  $\Phi_S$  we can compute:

$$\begin{aligned} \widetilde{Avg\_Price}_{kj} &\equiv \widetilde{\Xi}_k \Phi_k^{\delta_k(1-\sigma_k)} \int \int \left( z(\ell')^{\delta_k} \right)^{\Lambda_k} \widehat{\psi}_{kj}(z, \ell') dz d\ell' \\ &= \frac{Avg\_Price_{kj}}{P_C^{(1-\sigma_k)(1-\delta_k)\lambda_k}} \\ \widetilde{Avg\_Price}_{Cf} &\equiv \widetilde{\Xi}_C \Phi_C^{\delta_C(1-\sigma_C)} \int \int \left( z(\ell')^{\delta_C} \right)^{\Lambda_C} (\exp(d_F \times \mathcal{I}_C^x(z, \ell')))^{-\delta_C \sigma_C \Lambda_C} \widehat{\psi}_{Cf}(z, \ell') dz d\ell' \\ &= \frac{Avg\_Price_{Cf}}{P_C^{(1-\sigma_C)(1-\delta_C)\lambda_C}} \end{aligned}$$

At this point, we can compute the following variables, as functions of  $M_C$  and  $M_S$ :

$$N_{Ci} = \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} M_C \quad (\text{S.1})$$

$$N_{Si} = \frac{\omega_{Si}}{\varrho_{Si}^{exit}} M_S \quad (\text{S.2})$$

$$N_{Cf} = \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \quad (\text{S.3})$$

$$N_{Sf} = \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S \quad (\text{S.4})$$

$$M_{Ci} = \omega_{Ci} M_C$$

$$M_{Si} = \omega_{Si} M_S$$

$$M_{Cf} = \omega_{Cf} M_C$$

$$M_{Sf} = \omega_{Sf} M_S$$

**Firm-level expenditures with sector S goods (fixed operating costs, hiring costs, entry costs, fixed export costs)**

$$\begin{aligned}
E_S = & \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C (Avg\_Hiring\_Costs_{Cf} + \bar{c}_{Cf}) \\
& + \frac{\omega_{Ci} M_C}{\varrho_{Ci}^{exit}} (Avg\_Hiring\_Costs_{Ci} + \bar{c}_{Ci}) \\
& + \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S (Avg\_Hiring\_Costs_{Sf} + \bar{c}_{Sf}) \\
& + \frac{\omega_{Si} M_S}{\varrho_{Si}^{exit}} (Avg\_Hiring\_Costs_{Si} + \bar{c}_{Si}) \\
& + \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C Fraction\_Export_{Cf} f_x \\
& + M_C c_{e,C} \\
& + M_S c_{e,S}
\end{aligned}$$

Define  $c_C$ :

$$\begin{aligned}
c_C \equiv \frac{E_{S,C}}{M_C} = & \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} (Avg\_Hiring\_Costs_{Cf} + \bar{c}_{Cf}) \\
& + \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} (Avg\_Hiring\_Costs_{Ci} + \bar{c}_{Ci}) \\
& + \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} Fraction\_Export_{Cf} f_x \\
& + c_{e,C},
\end{aligned}$$

Where  $E_{S,C}$  is firm-level expenditures with sector  $S$  goods (fixed costs, etc) coming from  $C$ -sector activity.

Define  $c_S$ :

$$\begin{aligned}
c_S \equiv \frac{E_{S,S}}{M_S} = & \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} (Avg\_Hiring\_Costs_{Sf} + \bar{c}_{Sf}) \\
& + \frac{\omega_{Si}}{\varrho_{Si}^{exit}} (Avg\_Hiring\_Costs_{Si} + \bar{c}_{Si}) \\
& + c_{e,S},
\end{aligned}$$

Where  $E_{S,S}$  is firm-level expenditures with sector  $S$  goods (fixed costs, etc) coming from  $S$ -sector activity.

We can therefore write:

$$\begin{aligned}
E_S &= E_{S,C} + E_{S,S} \\
&= c_C M_C + c_S M_S
\end{aligned}$$

**Market Clearing ( $C$  and  $S$  sectors)**

Let  $I$  denote aggregate income. Then, market clearing in the  $C$  and  $S$  sectors must lead to:

$$\begin{aligned}\zeta I + X_C^{int} &= Rev_C - Exports + \tau_a Imports \\ (1 - \zeta) I + X_S^{int} + E_S &= Rev_S \\ Imports &= Exports\end{aligned}$$

Note that expenditures on intermediates are proportional to gross revenues:

$$P_k^m \iota_k(z, \ell) = \frac{(1 - \delta_k)(\sigma_k - 1)}{\sigma_k} R_k(z, \ell),$$

which leads to:

$$\begin{aligned}X_C^{int} &= \lambda_C \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C} Rev_C \\ &\quad + \lambda_S \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S} Rev_S \\ X_S^{int} &= (1 - \lambda_C) \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C} Rev_C \\ &\quad + (1 - \lambda_S) \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S} Rev_S\end{aligned}$$

Where  $Rev_C$  and  $Rev_S$  are total gross revenues in sectors  $C$  and  $S$  respectively. Therefore:

$$\begin{aligned}I &= \left(1 - \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C}\right) Rev_C \\ &\quad + \left(1 - \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S}\right) Rev_S \\ &\quad - E_S \\ &\quad + (\tau_a - 1) Exports\end{aligned}$$

Using

$$\begin{aligned}Rev_C &= Avg\_Revenue_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C + Avg\_Revenue_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} M_C \\ Rev_S &= Avg\_Revenue_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S + Avg\_Revenue_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} M_S \\ Exports &= Avg\_Exports_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \\ E_S &= c_C M_C + c_S M_S\end{aligned}$$

**Step 5a:** Write aggregate income  $I$  as a function of masses of entrants  $M_C$  and  $M_S$ .



$$\begin{aligned}
I = & \left(1 - \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C}\right) \left( \begin{aligned} & Avg\_Revenue_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \\ & + Avg\_Revenue_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} M_C \end{aligned} \right) \\
& + \left(1 - \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S}\right) \left( \begin{aligned} & Avg\_Revenues_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S \\ & + Avg\_Revenues_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} M_S \end{aligned} \right) \\
& - (c_C M_C + c_S M_S) \\
& + (\tau_a - 1) \left( Avg\_Exports_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \right)
\end{aligned}$$

Therefore:

$$I = a_C M_C + a_S M_S \quad (\text{S.5})$$

Where

$$\begin{aligned}
a_C = & \left(1 - \frac{(1 - \delta_C)(\sigma_C - 1)}{\sigma_C}\right) \left( \begin{aligned} & Avg\_Revenue_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} \\ & + Avg\_Revenue_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \end{aligned} \right) \\
& + (\tau_a - 1) \left( Avg\_Exports_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} \right) \\
& - c_C \\
a_S = & \left(1 - \frac{(1 - \delta_S)(\sigma_S - 1)}{\sigma_S}\right) \left( \begin{aligned} & Avg\_Revenues_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} \\ & + Avg\_Revenues_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \end{aligned} \right) \\
& - c_S
\end{aligned}$$

**Step 5b:** Write  $M_C$  as a functions of  $P_C$  and  $M_S$  as a function of  $M_C$  and  $\bar{P}_S$ .

**Price Index Sector C**

$$P_C^{1-\sigma_C} = P_{H,C}^{1-\sigma_C} + P_{F,C}^{1-\sigma_C}$$

The domestic component is given by:

$$\begin{aligned}
P_{C,H}^{1-\sigma_C} &= N_{Cf} Avg\_Price_{Cf} + N_{Ci} Avg\_Price_{Ci} \\
&= \left( \begin{aligned} & \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} Avg\_Price_{Cf} \\ & + \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} Avg\_Price_{Ci} \end{aligned} \right) M_C \\
&= \left( \begin{aligned} & \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} \widetilde{Avg\_Price}_{Cf} P_C^{(1-\sigma_C)(1-\delta_C)\lambda_C} \\ & + \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \widetilde{Avg\_Price}_{Ci} P_C^{(1-\sigma_C)(1-\delta_C)\lambda_C} \end{aligned} \right) M_C
\end{aligned}$$

We can therefore write  $P_{C,H}$  as:

$$P_{C,H}^{1-\sigma_C} = P_C^{(1-\sigma_C)(1-\delta_C)\lambda_C} b_C^1 M_C,$$

Where

$$b_C^1 \equiv \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} \widetilde{Avg\_Price}_{Cf} + \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \widetilde{Avg\_Price}_{Ci}$$

The foreign component is given by:

$$P_{F,C}^{1-\sigma_C} = (\epsilon \tau_a \tau_c)^{1-\sigma_C}.$$

Under Trade Balance:

$$\begin{aligned} Exports &= \frac{D_{H,C} (\epsilon \tau_a \tau_c)^{1-\sigma_C}}{\tau_a}, \\ \Rightarrow (\epsilon \tau_a \tau_c)^{1-\sigma_C} &= \frac{\tau_a \times Exports}{D_{H,C}} \\ &= \frac{\tau_a \times N_{Cf} Avg\_Exports_{Cf}}{D_{H,C}} \\ &= \frac{\tau_a \times Avg\_Exports_{Cf}}{\exp(\sigma_C \times d_{H,C})} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C \\ &= (P_C)^{\lambda_C(1-\delta_C)(1-\sigma_C)} \frac{\tau_a \times Avg\_Exports_{Cf}}{\Phi_C^{\frac{\sigma_C-1}{\lambda_C}}} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C. \end{aligned}$$

Where we have used

$$\exp(\sigma_C \times d_{H,C}) = \Phi_C^{\frac{\sigma_C-1}{\lambda_C}} (P_C)^{\lambda_C(1-\delta_C)(\sigma_C-1)}.$$

Therefore:

$$P_{F,C}^{1-\sigma_C} = (P_C)^{\lambda_C(1-\delta_C)(1-\sigma_C)} b_C^2 M_C,$$

Where

$$b_C^2 \equiv \frac{\tau_a \times Avg\_Exports_{Cf}}{\Phi_C^{\frac{\sigma_C-1}{\lambda_C}}} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}}.$$

Rewriting:

$$\begin{aligned} P_C^{1-\sigma_C} &= P_{C,H}^{1-\sigma_C} + P_{F,C}^{1-\sigma_C} \\ &= (P_C)^{\lambda_C(1-\delta_C)(1-\sigma_C)} b_C^1 M_C + (P_C)^{\lambda_C(1-\delta_C)(1-\sigma_C)} b_C^2 M_C \end{aligned}$$

So that:

$$P_C^{1-\sigma_C} = (b_C M_C)^{\frac{1}{(1-\lambda_C(1-\delta_C))}} \quad (\text{S.6})$$

where

$$b_C \equiv b_C^1 + b_C^2.$$

## Price Index Sector $S$

$$\begin{aligned}
P_S^{1-\sigma_S} &= N_{Sf} \text{Avg\_Price}_{Sf} + N_{Si} \text{Avg\_Price}_{Si} \\
&= \left( \frac{\frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} \text{Avg\_Price}_{Sf}}{\frac{\omega_{Si}}{\varrho_{Si}^{exit}} \text{Avg\_Price}_{Si}} \right) M_S \\
&= \left( \frac{\frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} \widetilde{\text{Avg\_Price}}_{Sf} P_C^{(1-\sigma_S)(1-\delta_S)\lambda_S}}{\frac{\omega_{Si}}{\varrho_{Si}^{exit}} \widetilde{\text{Avg\_Price}}_{Si} P_C^{(1-\sigma_S)(1-\delta_S)\lambda_S}} \right) M_S \\
&\Rightarrow P_S^{1-\sigma_S} = P_C^{(1-\sigma_S)(1-\delta_S)\lambda_S} b_S M_S
\end{aligned}$$

Where

$$b_S \equiv \frac{\frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} \widetilde{\text{Avg\_Price}}_{Sf} + \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \widetilde{\text{Avg\_Price}}_{Si}}$$

Given that  $P_S = \bar{P}_S$  is fixed, we can also write  $M_S$  as a function of  $P_C$  and model parameters.

$$M_S = \frac{\bar{P}_S^{1-\sigma_S}}{b_S P_C^{(1-\sigma_S)(1-\delta_S)\lambda_S}}, \quad (\text{S.7})$$

and using (S.6):

$$M_S = \frac{\bar{P}_S^{1-\sigma_S}}{b_S (b_C M_C)^{\frac{(1-\sigma_S)(1-\delta_S)\lambda_S}{(1-\sigma_C)(1-(1-\delta_C)\lambda_C)}}}. \quad (\text{S.8})$$

**Step 5c:** Write  $X_C^{int}$  as a function of  $M_C$  and  $M_S$ .

$$\begin{aligned}
X_C^{int} &= \lambda_C \frac{(1-\delta_C)(\sigma_C-1)}{\sigma_C} \left( \text{Avg\_Revenue}_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} + \text{Avg\_Revenue}_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \right) M_C \\
&\quad + \lambda_S \frac{(1-\delta_S)(\sigma_S-1)}{\sigma_S} \left( \text{Avg\_Revenue}_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} + \text{Avg\_Revenue}_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \right) M_S \\
&= d_C M_C + d_S M_S
\end{aligned}$$

Where

$$\begin{aligned}
d_C &= \lambda_C \frac{(1-\delta_C)(\sigma_C-1)}{\sigma_C} \left( \text{Avg\_Revenue}_{Cf} \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} + \text{Avg\_Revenue}_{Ci} \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} \right) \\
d_S &= \lambda_S \frac{(1-\delta_S)(\sigma_S-1)}{\sigma_S} \left( \text{Avg\_Revenue}_{Sf} \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} + \text{Avg\_Revenue}_{Si} \frac{\omega_{Si}}{\varrho_{Si}^{exit}} \right)
\end{aligned}$$

**Step 5d:** Pin down  $P_C$  using the equation defining  $\Phi_C$ , and obtain  $M_C$  and  $M_S$ .

We can now express aggregate income  $I$  as a function of  $P_C$  using equations (S.5), (S.6) and (S.7) and solve

for  $P_C$ . Remember that:

$$\exp(d_{H,C}) = \left( \frac{\zeta I + X_C^{int}}{P_C^{1-\sigma_C}} \right)^{\frac{1}{\sigma_C}}$$

Using the formula defining  $\Phi_C$  and manipulating, we obtain:

$$P_C^{-\lambda_C(1-\sigma_C)(1-\delta_C)} \Phi_C^{\frac{\sigma_C-1}{\lambda_C}} = \exp(\sigma_C \times d_{H,C}) = \frac{\zeta(a_C M_C + a_S M_S) + d_C M_C + d_S M_S}{P_C^{1-\sigma_C}},$$

which leads to

$$\Phi_C^{\frac{\sigma_C-1}{\lambda_C}} = \frac{(\zeta a_C + d_C) M_C + (\zeta a_S + d_S) M_S}{b_C M_C},$$

which allows us to solve for  $M_C$

$$M_C = \frac{1}{b_C} \left( \frac{(\zeta a_S + d_S) \bar{P}_S^{1-\sigma_S}}{b_S \left( \Phi_C^{\frac{\sigma_C-1}{\lambda_C}} - \frac{(\zeta a_C + d_C)}{b_C} \right)} \right)^{\frac{1}{1 + \frac{(1-\sigma_S)(1-\delta_S)\lambda_S}{(1-\sigma_C)(1-(1-\delta_C)\lambda_C)}}},$$

and then for  $M_S$  using (S.8).

**Step 5e:** Now that we have values of  $M_C$  and  $M_S$ , we obtain masses of firms  $N_{kj}$ .

$$N_{Ci} = \frac{\omega_{Ci}}{\varrho_{Ci}^{exit}} M_C$$

$$N_{Si} = \frac{\omega_{Si}}{\varrho_{Si}^{exit}} M_S$$

$$N_{Cf} = \frac{\varrho_{Ci}^{change} \omega_{Ci} + \varrho_{Ci}^{exit} \omega_{Cf}}{\varrho_{Cf}^{exit} \varrho_{Ci}^{exit}} M_C$$

$$N_{Cf} = \frac{\varrho_{Si}^{change} \omega_{Si} + \varrho_{Si}^{exit} \omega_{Sf}}{\varrho_{Sf}^{exit} \varrho_{Si}^{exit}} M_S$$

**Step 5f:** Obtain aggregate posted vacancies  $V_{kj}$ .

Now, substituting the expressions for the  $N_{kj}$ 's to obtain the number of vacancies in each sector as a function of  $L_u$ :

$$V_{Cf} = N_{Cf} Avg\_Vacancies_{Cf} + \frac{\omega_{Cf} M_C}{\mu^v}$$

$$V_{Ci} = N_{Ci} Avg\_Vacancies_{Ci} + \frac{\omega_{Ci} M_C}{\mu^v}$$

$$V_{Sf} = N_{Sf} Avg\_Vacancies_{Sf} + \frac{\omega_{Sf} M_S}{\mu^v}$$

$$V_{Si} = N_{Si} Avg\_Vacancies_{Si} + \frac{\omega_{Si} M_S}{\mu^v}$$

**Step 5g:** Save the values for  $P_C$  and  $P_{F,C}$ :

$$P_C^{1-\sigma_C} = (b_C M_C)^{\frac{1}{(1-(1-\delta_C)\lambda_C)}}$$

$$P_{F,C}^{1-\sigma_C} = (P_C)^{-\lambda_C(1-\delta_C)(\sigma_C-1)} b_C^2 M_C$$