

# Labour Market Monopsony Power and the Dynamic Gains to Openness Reforms\*

Priyaranjan Jha

*University of California, Irvine and CESifo*

Antonio Rodriguez-Lopez

*University of California, Irvine and CESifo*

Adam Hal Spencer

*University of Bonn*

This version: January 2026

## Abstract

We embed labour market monopsony into a dynamic heterogeneous-firm general equilibrium model with exporting, horizontal FDI, and rich firm lifecycle dynamics. Rising marginal costs with monopsony slow and limit incumbent firm growth in response to liberalisation, shifting adjustment to the extensive margin. Calibrated to U.S. micro data, welfare gains from tariff reduction are over four times larger under monopsony than with perfect labour markets. The difference is mostly driven by new exporter creation and firm entry along the transition path. By contrast, lowering outward FDI taxes gives powerful quantitative welfare losses under monopsony, as firms undertake FDI to escape domestic wage pressure.

**Keywords:** Firm life cycle; Monopsonistic labour market; Trade liberalisation; Dynamics; Foreign direct investment

**JEL codes:** F12, F13, F16, F23, F40, D25

---

\*We are deeply grateful to Dirk Krueger, Michael Waugh and three anonymous referees for comments that have significantly improved the paper. We also thank Andres Rodriguez-Clare, Christian Pröbsting, Fernando Parro, George Alessandria, Kim Ruhl, Paweł Strawiński and Jaume Ventura for greatly insightful comments. We also are thankful for participation in presentations at Michigan State University, Federal Reserve Board of Governors, University of Bonn, Bank of Lithuania, Universitat Pompeu Fabra, University of Kent, Banque de France, University of Cagliari, EIEF, University of Naples Federico II, University of Mannheim, Bocconi University, Liverpool Macro Workshop, Workshop of the Spanish Macro Network, Workshop on Frontiers of Macro-Labour, 2023 Lisbon Macroeconomics Workshop (Católica Lisbon School of Business and Economics), 2023 Australasian Meeting of the Econometric Society (University of New South Wales), 5<sup>th</sup> Warsaw Money-Macro-Finance Conference (University of Warsaw), 2023 Fall Midwest Macroeconomics Meeting (Texas Tech University), 2023 European Trade Study Group Meeting (ETSG, University of Surrey), and 2025 Western Economic Association International Meeting (San Francisco). All errors are our own. Corresponding author: Adam Hal Spencer ([spencer@uni-bonn.de](mailto:spencer@uni-bonn.de)).

# 1 Introduction

Growing empirical evidence suggests that firms face upward-sloping labour supply curves — they operate in a monopsonistic environment (Naidu et al., 2018; Yeh et al., 2022). This stands in contrast with the standard assumption in the trade literature that firm-level labour supply is perfectly elastic.<sup>1</sup> We ask how labour market monopsony affects the firm lifecycle and what this means for the short-run and long-run gains from openness reforms. A developing literature incorporates monopsony into theoretical trade models (MacKenzie, 2019; Jha and Rodriguez-Lopez, 2021; Egger et al., 2022; Dhyne et al., 2022). We contribute to this literature by building a quantitative laboratory of firm dynamics, trade, and FDI, which can be disciplined with micro data and solved for both the steady state and the transition path, following a policy action.

World trade policy has shifted abruptly in recent years, with effective tariff rates in major economies moving sharply and at times unpredictably. Such rapid and uncertain changes underscore the importance of understanding short-run adjustment dynamics. Firms cannot instantaneously alter their foreign engagement: new exporters and multinational firms face substantial fixed costs and expand only gradually following entry (Gumpert et al., 2020). Capturing monopsony in a dynamic setting is therefore essential, both to interpret the effects of large policy shocks as they unfold over time and to quantify welfare accurately.<sup>2</sup>

Our model embeds a rich environment for firm cost structures, where monopsony (often referred to as an *imperfect labour market* structure) exerts a first-order influence on firm behaviour. When firms face upward-sloping labour supply (USLS), consistent with the empirical evidence on monopsony discussed below, their marginal cost of production rises with scale. That is, hiring an additional worker requires increasing the wage for all existing employees. This contrasts with the perfectly elastic case (often referred to as a *perfect labour market* structure), where marginal cost is flat regardless of size. As a result, rising marginal costs can constrain firm growth, limiting their ability to scale in response to the profit opportunities created by liberalisation. These effects spill over to younger, domestic firms, since foreign engagement typically occurs later in the firm lifecycle. Aggregate adjustment thus shifts away from the intensive towards the extensive margin of foreign engagement — a central force behind the results that follow.

To model the firm lifecycle, from entry to exporting to multinational activity, we embed monopsonistic labour markets in an otherwise standard dynamic general equilibrium framework with stochastic time-varying firm-level states and endogenous entry and exit. Firms are monopsonistically competitive in the labour market, with labour being their only factor of production. They behave monopolistically competitively in goods markets and face an idiosyncratic fundamental productivity shock, which is a persistent process that evolves over time. Each period, firms make a discrete choice — whether to exit the industry, operate as a domestic firm, an exporter, or a horizontal multinational — that depends on their individual state. Fixed cost structures follow the approach in Alessandria and Choi (2007) and

---

<sup>1</sup>That is, firms can hire as much labour as they desire at the fixed market wage rate — labour supply in the aggregate can still be variable.

<sup>2</sup>Using static trade models or comparing steady states alone can lead to misleading normative conclusions (see Alessandria et al., 2021b and Spencer, 2022).

Alessandria et al. (2021b), with a one-time sunk cost to establish an operating segment and fixed continuation costs thereafter.<sup>3</sup> The exporting–FDI margin builds on Helpman, Melitz and Yeaple (2004) and Arkolakis et al. (2018), incorporating variable inefficiencies (iceberg costs) on top of these fixed costs of foreign engagement. These iceberg costs vary over time as in Alessandria et al. (2021b) and are specific to the firm’s current mode of production (exporting or multinational). There are two values of each mode-specific iceberg cost; when opening a new export branch or foreign subsidiary, firms start with a high cost, which transits to a lower value over time at a stochastic rate. This structure captures the gradual growth patterns of exporters and multinationals following initial investment. We discipline these costs using micro data on the cross-section of firms. Armed with the calibrated model, we then study the effect of two sets of liberalisation episodes — tariff and FDI tax reforms — that operate on the top end of the firm lifecycle.

We study the implications of trade liberalisation under monopsonistic labour markets through a counterfactual that bilaterally removes tariffs, starting from a 10 percent benchmark.<sup>4</sup> We compare the resulting transition and steady state to a perfect labour market parameterisation, where the initial tariff is calibrated to equalise long-run trade elasticities across the two environments. The key insight from this quantitative exercise is that the welfare gains from bilateral trade liberalisation — measured in lifetime consumption-equivalent variation — are more than four times greater under monopsony (2.03%) than with perfect labour markets (0.44%). However, along the transition path, the short-run consumption losses are significantly larger under monopsony, with an impact drop of consumption of around 5%, contrasting with a weaker contraction of around 1.5% with perfect markets. Such a contrast can be important in settings where short-run adjustment costs are given disproportionate weight in policy decisions.

Contrasting trade liberalisation predictions arise from how monopsony reshapes firm-level responses to tariff cuts. Under monopsonistic labour markets, rising marginal costs make aggressive expansion unattractive for incumbent exporters — intensive margin adjustment is weaker. Instead, trade liberalisation triggers a strong response along the extensive margin — many firms switch into export status and new firms enter along the transition to exploit improved profitability. Although total firm mass declines in the new steady state — as in standard trade models — the exporter mass rises sharply and stabilises quickly. Because incumbents are constrained from scaling up, competitive pressure remains moderate, increasing the survival rates of entrants and new exporters. The result is a smoother and faster transition to the new steady state. In contrast, perfect labour markets enable rapid incumbent expansion, crowding out entrants at all points along the transition and raising exit rates as competition intensifies. Under imperfect labour markets, the resulting variety expansion — driven by a larger exporter base and a milder decline in total firm mass — translates into significantly stronger welfare gains (through households’ love of variety in consumption). Note, however, that the creation of these new varieties and export branches along the transition requires a larger equity injection into firms from households, driving the

---

<sup>3</sup>See also Alessandria and Choi (2014a), Alessandria and Choi (2014b), and Ruhl and Willis (2017) for related treatments of fixed cost structures.

<sup>4</sup>We focus on bilateral reforms under country symmetry to keep the discussion of mechanisms clear and tractable. This choice reflects the reciprocal and retaliatory nature of many trade and investment agreements.

stronger short-run consumption fall with imperfect markets.

Our finding that welfare gains from tariff liberalisation are larger under imperfect labour markets is fundamentally a quantitative result. Analytically, the model shows that monopsony weakens the intensive margin trade elasticity and strengthens that of the extensive margin, but it does not deliver an analytical ordering of welfare changes across labour market structures. In the quantitative exercise, we recalibrate fixed and variable costs in each environment so that the initial steady state and the long-run trade elasticity are held constant across labour market structures. We then decompose the implied elasticity into intensive and extensive components — the extensive margin accounts for 72% of the total trade elasticity under monopsony, compared with 67% under perfect labour markets. It is this calibrated shift in the composition of adjustment towards exporter entry that generates a stronger expansion in the mass of varieties entering the CES aggregator, thereby delivering larger welfare gains.

Having explored how monopsony shapes the domestic–exporter decision, we then turn to monopsony’s impact on the later stage of a firm’s foreign engagement lifecycle — multinational activity. Monopsony gives rise to what we term the *escape effect*. When a firm transitions from domestic to exporter status, production for both local and foreign sales relies on hiring from a single local labour supply function, putting upward pressure on wages and raising marginal costs. By becoming a multinational, the firm escapes this constraint: it can serve the domestic market from headquarters while producing foreign sales through a separate subsidiary.<sup>5</sup>

Under monopsony, the escape effect makes the benefits of becoming a multinational *front-loaded*: an upgrading exporter immediately lowers its marginal costs, making the higher fixed and initial FDI-specific iceberg costs worthwhile. As a result, many exporters with low trade costs upgrade to multinational status. This contrasts with the perfect labour market case, where marginal costs are independent of size — most firms that upgrade are those with high export costs seeking to realise future gains. Thus, under a perfect labour market, the benefits of multinational status are instead *back-loaded*: firms upgrade despite high initial costs, hoping to quickly draw a low FDI iceberg cost in subsequent periods.

While the option to escape is valuable for profits from an individual firm’s perspective, it leads to lower wages and an inefficiency whereby more FDI can be harmful from a social perspective — a result we show analytically using a simple model. We find that this effect is large quantitatively, in the context of an outward FDI tax liberalisation experiment.<sup>6</sup> We find that a 1% multinational profit tax liberalisation is roughly neutral in terms of welfare with perfect labour markets, while it gives lifetime welfare losses of 0.5% with monopsony. Importantly, these effects can only be quantified in a dynamic framework that disciplines the firm lifecycle, since FDI is a rare event that typically occurs late in a firm’s life.

---

<sup>5</sup>In both the parameterisation and the data, firms almost always become multinationals through first operating as an exporter.

<sup>6</sup>We consider outward FDI taxes as the instrument for FDI liberalisation, motivated by recent policy discussions around taxation and international investment. However, the modelling of this instrument is stylised, and we view this exercise as illustrative of mechanisms rather than as a structural evaluation of actual tax reforms.

Our quantitative results reveal adjustment patterns that move in unexpected directions. Monopsony weakens intensive margin responses but amplifies extensive margin adjustment; it generates larger welfare gains from trade liberalisation despite weaker firm selection. At the same time, it makes outward FDI potentially welfare-reducing, even though FDI is conventionally viewed as a sign of competitiveness and openness. These counterintuitive outcomes highlight the importance of accounting for labour market frictions and firm dynamics when undertaking policy evaluation exercises regarding the open economy.

The remainder of the paper is structured as follows. Section 2 provides a review of the relevant literature. Section 3 outlines the general environment of our dynamic model. Section 4 takes a static special case of this environment to obtain an analytical characterisation of the equilibrium. We include this section with a view towards developing a strong intuition and understanding of the inefficiencies present in the model, as well as the role of its distinctive features. Section 5 then returns to the general infinite-horizon model and gives the Bellman equations for firms’ dynamic decisions, as well as the full equilibrium definition. Section 6 details the parameterisation strategy. Section 7 presents the baseline quantitative results, while Section 8 gives parameter robustness and extensions. Finally Section 9 concludes.

## 2 Related Literature

A rapidly growing strand of the trade literature embeds upward-sloping firm-level labour supply curves — and hence wage-markdown power — into otherwise standard models of heterogeneous firms. A common insight is that monopsony alters how firms transmit international shocks, although the mechanisms emphasised vary across studies.

MacKenzie (2019) constructs a static model where upward-sloping labour supply curves arise from idiosyncratic productivity matches between firms and workers. The firms face oligopoly in the product market with endogenous markups and oligopsony in the factor market with endogenous markdowns. In quantitative analysis using data from India, he finds that trade liberalisation has a larger effect on markups in the product market than on markdowns in the labour market. In our framework, the markups and markdowns are constant but the mass of firms is endogenous and much of the impact of trade liberalisation occurs through the extensive margin. Jha and Rodriguez-Lopez (2021) also rely on idiosyncratic preferences, but show that monopsony can amplify or dampen welfare gains from goods trade liberalisation through a novel “love for employer variety” channel. Whereas Jha and Rodriguez-Lopez (2021) present a qualitative, steady-state analysis (deriving results under a Pareto productivity distribution), we conduct a quantitative evaluation of trade liberalisation that considers welfare changes both in steady state and along the transition path. In implementing this quantification, we recalibrate key fixed and variable cost parameters in each scenario (monopsonistic vs. perfect labour market) so that certain empirical moments remain consistent; by contrast, Jha and Rodriguez-Lopez (2021) hold those parameters fixed across their labour market scenarios. This calibrated quantitative approach allows us to isolate the impact of labour market structure on the results more cleanly.

An important and related contribution by Dhyne, Kikkawa, Komatsu, Mogstad and Tintelnot (2022) examines the labour-market effects of foreign demand (trade) shocks in a model

that embeds monopsony power within a production network and incorporates fixed overhead costs. They find that combining monopsonistic labour markets with fixed costs amplifies the impact of trade shocks relative to a competitive-labour benchmark without fixed costs, with the fixed-cost channel playing the primary quantitative role. Our framework complements theirs along two dimensions. First, we abstract from the production network but introduce firm entry and exit, which creates an extensive-margin adjustment channel under trade liberalisation. We show that, with this extensive margin, monopsony power alone can serve as a quantitatively important transmission mechanism for trade shocks, even holding the fixed cost environment constant. Second, whereas their work highlights network propagation of foreign demand shocks, we focus on how monopsony alters individual firm life-cycle behaviour and the path of aggregate adjustment following liberalisation.

Notably, unlike all the above studies, our model also allows firms to undertake horizontal foreign direct investment (FDI), and we explicitly examine the implications of FDI liberalisation. [Egger et al. \(2022\)](#) provide a related perspective by incorporating a monopsonistic labour market in a model of trade with heterogeneous firms. They do not allow firms to engage in horizontal FDI as in our current model, but do allow firms to offshore inputs from abroad. Just as an upward-sloping labour supply curve incentivises firms to engage in horizontal FDI compared to exporting in our setup, in their framework firms are induced to offshore input production. They find that trade in goods is welfare improving but offshoring can be welfare reducing because it allows firms to exercise their labour market power by reducing the size of their domestic operations. Our result on the negative welfare implications of a cut in FDI taxes qualitatively echoes their finding of adverse welfare effects of offshoring. We push the analysis further by quantifying the size of this effect using a dynamic model that can be disciplined with firm micro data.

More broadly, our work connects to the large literature on firm dynamics in open economies. We focus here on those that are close to us in terms of the multinational choice of firms.<sup>7</sup> [Spencer \(2022\)](#) studies the implications of corporate tax reforms targeted at multinational firms in the presence of corporate finance frictions but perfect labour markets. [Gumpert, Li, Moxnes, Ramondo and Tintelnot \(2020\)](#) provide empirical details regarding the exporter and multinational lifecycle and incorporate such choices into a dynamic discrete choice framework, but abstract from modelling the intensive margin. [Garetto, Oldenski and Ramondo \(2019\)](#) study a rich multi-country structure with export platforms and dynamic FDI choices in the context of perfect labour markets. We contribute to this literature by quantifying the option to escape local upward-sloping labour supply curves and how it affects aggregate dynamics.

Our analysis also relates to research on labour market frictions in macroeconomics and labour economics. Our theoretical model shares some features of the [Berger, Herkenhoff and Mongey \(2022\)](#) model, where labour supply is endogenous. Their focus is on estimating and quantifying the welfare effects of labour market power on the levels of aggregate variables in a closed-economy setting. They find a significant welfare loss from labour market power

---

<sup>7</sup>See [Alessandria, Arkolakis and Ruhl \(2021a\)](#) for a summary of the literature on firm dynamics and trade. See [Spencer \(Forthcoming\)](#) for a comprehensive review of the firm dynamics and multinational activity literature.

compared to the efficient allocation. In a related paper on minimum wage policy, [Berger, Herkenhoff and Mongey \(2025\)](#) show that when firms possess wage setting power, a modest minimum wage increase can raise aggregate welfare, even after accounting for potential employment losses, because it compresses wage markdowns and reallocates labour towards more productive firms. [Jha and Rodriguez-Lopez \(2022\)](#) also use a similar model to study the effects of minimum wages when workers care about employer variety. Their theoretical model shows that the welfare effects of minimum wages may be negative due to changes in the mass of employers.

Finally, a substantial body of empirical research provides evidence for finitely elastic (and hence upward-sloping) labour supply at the firm level. [Naidu et al. \(2018\)](#) survey the literature and provide a range of 1 to 5 for the labour supply elasticities that firms face. [Berger, Herkenhoff and Mongey \(2022\)](#) estimate firm-level labour supply elasticities ranging from 0.76 (for a firm that controls an entire local labour market) to 3.74 for the smallest firms. [Webber \(2015\)](#) uses the U.S. Census’s Longitudinal Employer-Household Dynamics (LEHD) data and obtains a mean labour supply elasticity of 1.08, whereas [Yeh, Macaluso and Herstein \(2022\)](#) find a mean value of 1.88; the latter is the number we use in our quantitative exercise. A consequence of a finitely elastic labour supply is that firms have market power and they markdown wages below the marginal product of labour. [Yeh, Macaluso and Herstein \(2022\)](#) estimate the markdowns directly using the output elasticity and revenue share of labour. They find the markdowns to be substantial and increasing since the early 2000s. Our quantitative exercises show that labour market power can significantly affect the welfare effects of trade and FDI reforms.

### 3 Model Environment

Our model consists of two symmetric countries, referred to as Home ( $H$ ) and Foreign ( $F$ ). There are four types of agents in each country: households, intermediate producers, final producers and governments. Labour in each country is the only variable input into production; this factor is immobile across countries. There is trade in intermediate-good varieties across countries, as well as horizontal FDI. We adopt the notation convention that variables with a \* superscript correspond to activities in  $F$ , while those without correspond to those in  $H$ . In the exposition that follows, we focus on agents from the  $H$  economy; the setup for those from  $F$  can be defined similarly. Aggregate and firm-level variables are denoted with capital and lower-case letters, respectively. Time is discrete and indexed by subscript  $t$ .

#### 3.1 Households and Monopsonistic Labour Markets

We assume a mass  $\bar{L}$  of households/workers. Each period, they are ex-ante homogeneous and supply labour to intermediate goods firms operating in their local economy. Producers of intermediate goods varieties are indexed by  $\omega$ , and the mass of producing firms is given by

$$\Omega_t^P = \Omega_t + \Omega_t^{M*},$$

where the  $P$  superscript in  $\Omega_t^P$  stands for *producing*. Variable  $\Omega_t$  is the mass of operative firms headquartered/incorporated in  $H$  (we will often refer to this as the measure of HQ firms), while  $\Omega_t^{M*}$  is the mass of multinational firms headquartered in  $F$ . Thus, households in  $H$  supply labour to both domestic firms and to foreign multinationals' local affiliates.

Each period, workers draw match-specific productivity levels  $\varepsilon_t(\omega)$  — or efficiency units of labour — from a Fréchet distribution with shape parameter  $\theta + 1$ .<sup>8</sup> Given a draw  $\varepsilon_t(\omega)$  and hours worked  $h_t(\omega)$ , the utility derived from employment at firm  $\omega$  is

$$U_t(\omega, h_t(\omega)) = \frac{\varepsilon_t(\omega)h_t(\omega)w_t(\omega) + (\Pi_t + T_t)/\bar{L}}{P_t} - \frac{\phi}{1 + \phi} [Ch_t(\omega)]^{\frac{1+\phi}{\phi}}, \quad (1)$$

where  $\phi < \theta$  is the Frisch elasticity of labour supply,  $P_t$  is the consumer price index (CPI), and  $C$  is a normalisation constant set equal to  $(\Omega_t^P)^{\frac{1}{1+\theta}}$  to neutralise the love of variety for employers.<sup>9</sup>

The parameter  $\theta$  governs the degree of monopsony power employers hold in the economy. Lower values of  $\theta$  imply greater dispersion in worker-firm match quality, yielding employers more monopsony power and giving scope for a non-degenerate wage distribution with a positive supply of workers to any given firm. Put differently, due to variation in relationship-specific productivity levels,  $\varepsilon_t(\omega)$ , not all workers will leave a given firm when the associated wage  $w_t(\omega)$  drops. As  $\theta \rightarrow \infty$ , all match-specific productivity draws centre around 1, giving the standard case in the literature where employers are homogeneous from a worker's perspective — the perfect labour market case.

In addition to utility derived from labour income, workers receive a per-capita share of firm profits net of fixed costs,  $\Pi_t/\bar{L}$ , and lump-sum tax rebates from the government  $T_t/\bar{L}$ . We assume local workers own the equity of all firms headquartered in their country of residence.

A worker chooses the amount of labour supplied to producer of variety  $\omega$ ,  $h_t(\omega)$ , in order to maximise their welfare of  $U_t(\omega, h_t(\omega))$  given in Equation (1). Doing so gives the solution

$$h_t(\omega) = (\Omega_t^P)^{-\frac{1+\phi}{1+\theta}} \left[ \frac{\varepsilon_t(\omega)w_t(\omega)}{P_t} \right]^\phi. \quad (2)$$

One can then use solution (2) in Equation (1) to get household utility from working for producer  $\omega$ , defined as  $U_t(\omega)$ . A worker will choose producer of variety  $\omega$  over  $\omega'$  when  $U_t(\omega) > U_t(\omega')$ . In Appendix A.1, we obtain the following expression for the labour supply to firm  $\omega$ :

$$n_t(\omega) = B_t w_t(\omega)^\theta, \quad (3)$$

---

<sup>8</sup>Formally,  $\Omega_t^P$  is a measurable set and  $\omega \in \Omega_t^P$  indexes a continuum of choices. The idiosyncratic shock  $\varepsilon_t(\omega)$  is a random variable defined for each  $\omega \in \Omega_t^P$ .

<sup>9</sup>In this specification, workers are risk neutral; we consider an extension with risk aversion in Section 8.3, and our long-run policy inferences are robust to this assumption. Additionally, the model exhibits a love-of-variety effect on the labour supply side — analogous to consumption-side love-of-variety in the random utility framework of Anderson, De Palma and Thisse (1992) — since a larger set of employers increases the probability of a high match-specific productivity draw. Because the empirical magnitude of this effect is difficult to identify, we neutralise it in our quantitative exercises.

where  $B_t \equiv \Gamma\left(\frac{\theta-\phi}{1+\theta}\right) \bar{L}/(P_t^\phi \Omega_t^P W_t^{\theta-\phi})$  is a supply-shifter,  $\Gamma$  is the gamma function, and  $W_t$  is the wage index,

$$W_t = (\Omega_t^P)^{-\frac{1}{1+\theta}} \left( \int_{\omega \in \Omega_t^P} w_t(\omega)^{1+\theta} d\omega \right)^{\frac{1}{1+\theta}}, \quad (4)$$

which represents the average wage in the economy.<sup>10</sup> From Equation (3), the labour supplied to the firm of variety  $\omega$  depends on a shifter-term comprised of aggregates, and a term that is increasing in the wage the firm offers. This is the upward-sloping labour supply curve (USLS, hereafter) that intermediate-good firms face.

One can then write (see Appendix A.1) household-level period  $t$  welfare (denoted  $EU_t$  for expected utility over employers) as

$$EU_t = \frac{1}{1+\phi} \left( \frac{W_t}{P_t} \right)^{1+\phi} \Gamma + \frac{(\Pi_t + T_t)/\bar{L}}{P_t}, \quad (5)$$

which gives expected utility as a transformation of the real wage in addition to real proceeds from profits and government re-distributions. The full lifetime utility can then be written as

$$EU = \sum_{t=0}^{\infty} \beta^t EU_t,$$

where  $\beta \in [0, 1]$  is the household's rate of time preference, which will be used in the firms' dynamic investment problem described below.

### 3.2 Final Goods Firms

A representative final goods producer aggregates over the intermediate goods varieties that are utilised domestically. We define the mass of varieties utilised as

$$\Omega_t^U = \Omega_t + \Omega_t^{X*} + \Omega_t^{M*},$$

where the  $U$  in  $\Omega_t^U$  stands for *utilised*,  $\Omega_t^{X*}$  is the mass of exporters sending goods from  $F$  to  $H$  and recall  $\Omega_t^{M*}$  is the mass of multinationals headquartered in  $F$  selling through their subsidiary in  $H$ . Final goods producers use technology

$$A_t = \left( \int_{\omega \in \Omega_t^U} q_t(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \quad (6)$$

where  $q_t(\omega)$  denotes the input of intermediate goods of variety  $\omega$  and  $\sigma > 1$  is the elasticity of substitution across varieties.

As in [Alessandria, Choi and Ruhl \(2021b\)](#), we introduce a parameter  $\hat{\tau}^X \geq 1$  as a tariff levied by the  $H$  government on imported goods from  $F$ . Given the symmetric structure, it

---

<sup>10</sup>In the remainder of the paper we suppress the argument  $\frac{\theta-\phi}{1+\theta}$  of the gamma function to reduce notational clutter.

also equals the tariff levied by  $F$  on goods imported from  $H$ . This tariff is borne by the final goods firms at the point they purchase imported varieties from abroad. These competitive final goods producers solve a profit maximisation problem that gives the following demand curves

$$\begin{aligned} q_t(\omega) &= A_t P_t^\sigma p_t(\omega)^{-\sigma} & \forall \omega \in \Omega_t + \Omega_t^{M*} \\ q_t(\omega) &= A_t P_t^\sigma [\widehat{\tau}^X p_t(\omega)]^{-\sigma} & \forall \omega \in \Omega_t^{*X}. \end{aligned} \quad (7)$$

Note that the final goods producer internalises the effect of the tariff as an increase in the variety's effective price. Finally, we define the CPI as

$$P_t = \left( \int_{\omega \in \Omega_t + \Omega_t^M} p_t(\omega)^{1-\sigma} d\omega + \int_{\omega \in \Omega_t^{*X}} (\widehat{\tau}^X)^{1-\sigma} p_t(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}, \quad (8)$$

where we take the CPI in  $H$  to be the numeraire.

### 3.3 Intermediate Goods Firms

Intermediate firms operate under a dynamic [Helpman, Melitz and Yeaple \(2004\)](#) environment with fixed and sunk costs for maintaining and assuming statuses, respectively. They are also subject to stochastic iceberg cost draws, as in [Alessandria, Choi and Ruhl \(2021b\)](#), for the purpose of matching the exporter and multinational lifecycles. An incumbent's state variable will change over time and it will always have the option to exit the industry. As such, we describe the environment for incumbent firms and new entrants in turn. To economise on notation, we omit variety-level notation in what follows.

#### 3.3.1 Incumbents

Incumbent firms maximise the discounted value of expected dividends they pay to their shareholders

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t d_t, \quad (9)$$

where  $\beta \in [0, 1]$  is the household's discount factor,  $d_t$  is the dividend in period  $t$ , and the expectation operator is with respect to their idiosyncratic shocks. These firms produce using a constant returns to scale technology

$$y_t = z_t n_t, \quad (10)$$

where  $y_t$  is output,  $z_t$  is the firm's idiosyncratic productivity level, and  $n_t$  is its employment of labour. The productivity level evolves over time through process

$$\log(z_t) = \rho_z \log(z_{t-1}) + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_z^2), \quad (11)$$

where  $\rho_z \in [0, 1)$  is a persistence parameter,  $\epsilon_t$  is a shock with normal distribution about a zero mean with variance  $\sigma_z^2$ . As shorthand, we refer to the conditional process defined by

(11) as  $g(z_t|z_{t-1})$ . We assume, to keep the state space small, that the same productivity process applies to production in  $F$  should this  $H$  firm choose to become a multinational. Time-varying productivity in (11) is what contributes to nontrivial transition dynamics in our model following a reform.

Firms make a choice of how much labour to hire, subject to the optimal labour supply choices of households in (3) and the intermediate-good demand curves in (7). A firm with operations in  $H$  must offer a wage to induce the desired supply of labour from the  $H$  household. A multinational will make the same considerations when making their labour choice in  $F$ .

**Extensive margin.** Incumbents make a choice about their operations at the extensive margin. They choose a status ( $s_{t+1}$ ) each period, which can be either to exit ( $E$ ), operate as a domestic firm ( $D$ ), an exporter ( $X$ ) or a multinational ( $M$ ).<sup>11</sup> A domestic firm sells its goods only to the market of its country of incorporation/headquarters. An exporter also sells to the final-good producer abroad, from its domestic production. A multinational establishes a local subsidiary abroad, employing in the local workforce, for production of goods sold in the local market.

As in papers such as [Alessandria and Choi \(2007\)](#), [Alessandria, Choi and Ruhl \(2021b\)](#) and [Ruhl and Willis \(2017\)](#), firms pay a one-off sunk cost to establish a new operating segment. They then pay fixed period-by-period costs to maintain it (and receive new productivity shocks) thereafter. Post-entry, all firms incur fixed cost  $f^C$  (the  $C$  is shorthand for *continuation*) each period they remain active. A firm changing its status to  $s_{t+1} \in \{X, M\}$  from status  $s_t \in \{D, X, M\}$  with  $s_{t+1} \neq s_t$ , incurs sunk cost  $f_t^{s_{t+1}, s_t}$ . That is, the sunk costs are allowed to vary with the firm's current status. A firm that maintains a status of  $s_{t+1} \in \{X, M\}$ , having incurred the associated sunk cost previously ( $s_{t+1} = s_t$ ), pays period-by-period cost  $f^{s_{t+1}, C}$ . Notice also the time subscript on  $f_t^{s_{t+1}, s_t}$ ; for technical reasons, we allow these four possible sunk costs ( $f_t^{X,D}, f_t^{M,D}, f_t^{X,M}, f_t^{M,X}$ ) to exhibit some randomness to generate dispersion. We then denote the vector of a firm's sunk costs as  $f_t$  and the probability distribution they follow by  $H(f_t)$ , which is iid across periods.<sup>12</sup> All of these fixed costs are denoted in terms of goods, but indexed to the aggregate wage index of the country in which they are headquartered. We give a detailed discussion of this assumption in [Section 3.5](#).

**Intensive margin.** Firms that export or undertake FDI are also subject to variable cost inefficiencies. The interpretation in the latter case is as a productivity loss associated with the operations of a foreign subsidiary (as in [Arkolakis, Ramondo, Rodríguez-Clare and Yeaple,](#)

<sup>11</sup>Note that although firms make the choice of  $s_t$  each period, the calibration of the fixed costs will be done to match the persistence of statuses in the data.

<sup>12</sup>Given that we solve this model using the recursive problem, our computations require discretising the firm state space. This sunk cost randomness gives more smoothness in firm choices and therefore better convergence properties for the numerical exercises, in particular with the transition path solution. Such is standard practice in models of firm dynamics (e.g. see [Sedláček, 2020](#); [Bendetti-Fasil, Impullitti, Licandro, Sedláček and Spencer, 2024](#)). In practice, this amounts to introducing another parameter  $\sigma_f$  that must be disciplined with data. Specifically, we assume that  $f_t^{s_{t+1}, s_t} \sim LN(\hat{f}_t^{s_{t+1}, s_t}, \sigma_f^2)$  where  $\hat{f}_t^{s_{t+1}, s_t}$  is the mean and  $\sigma_f^2$  is the variance of a log-normal distribution for  $s_{t+1} \in \{X, M\}$  and  $s_t \in \{D, X, M\}$  with  $s_{t+1} \neq s_t$ . These costs are drawn iid over time. Note that  $\sigma_f^2$  is the same for all four sunk costs.

2018). Following Alessandria, Choi and Ruhl (2021b), we assume that firms with domestic status  $D$  face infinite variable foreign engagement costs,  $\tau^D = \infty$ . However, a firm upgrading to status  $s \in \{X, M\}$  will start its new branch with an iceberg cost of  $\tau_H^s > 1$ , where the subscript  $H$  denotes “high.” That is, payment of the sunk cost for  $s_{t+1} \in \{X, M\}$  of  $f_t^{s_{t+1}, s_t}$  allows a firm to serve the  $F$  market through status  $s_{t+1}$  with variable inefficiency  $\tau_H^{s_{t+1}}$ . Thereafter, maintaining this status through payment of the fixed period-by-period cost  $f^{s,C}$  yields the firm a new draw of this iceberg cost from the set  $\{\tau_L^s, \tau_H^s\}$  where  $1 \leq \tau_L^s < \tau_H^s$ ; the subscript  $L$  denotes “low.” The iceberg cost follows a Markov process with persistence probability of  $0 < \rho_\tau^s < 1$  for  $s \in \{X, M\}$ . This process can be interpreted as firms investing over time to lower their variable foreign engagement costs: firms with  $\tau_H^s$  ( $\tau_L^s$ ) are typically younger (older) in their foreign-production mode, on average.<sup>13</sup> As such, we will often refer to firms with  $\tau_H^s$  as being young in  $s$  activity, while those with  $\tau_L^s$  as mature. Formally, we denote this process by  $g_\tau(\tau_{t+1}^{s_{t+1}}, s_{t+1} | \tau_t^{s_t}, s_t)$ .

Incumbent firms who generate profits through multinational activity are taxed by the government of their country of incorporation on such profits at rate  $\hat{\tau}^M \in [0, 1]$ . Note the difference in the treatment of the distortions levied by the government across imports and FDI. The tariff rate is on the sale value of imported goods, while the FDI tax is levied on corporate profits of multinationals abroad. We follow this approach to approximate the type of policies used by governments in practice, in relation to firms of each status.<sup>14</sup> The tariff distorts the exporting firm’s demand curve, potentially affecting both the extensive and intensive margins. This differs from the FDI tax, which solely affects the extensive margin choice, given that labour expenses are tax deductible.

**Market structure.** We make the standard assumption of monopolistic competition on the goods side, with firms setting constant markups. We then make the analogous assumption on the input side: firms engage in monopsonistic competition in the labour market, setting a constant markdown. The elasticity parameters in firm demand and labour supply determine the extent of market power, with markups given by  $\sigma/(\sigma - 1)$  and markdowns by  $\theta/(1 + \theta)$ . As such, we study the quantitative effect of monopsony on firms’ cost structures and how that aggregates to affect the masses of firms and the macroeconomy.

**Timing.** An incumbent faces the following timing.

1. Enter the period with state vector  $(z_{t-1}, s_t, \tau_t^{s_t})$ .
2. Draw new productivity shock  $z_t$  from distribution in (11).
3. Make static production choices in accordance with  $s_t$ ,  $\tau_t^{s_t}$  and newly-drawn  $z_t$ .
4. Draw sunk cost shocks.

---

<sup>13</sup>Note in the absence of stochastic variable costs, that is when all firms have the same iceberg costs, the model gives the counterfactual prediction that new exporters/multinationals are more productive than those continuing. It also predicts the absence of growth in foreign engagement intensity — all firms jump immediately to their long run intensity in the first year of operation.

<sup>14</sup>The outward tax on FDI can be thought of as resembling the repatriation tax levied by the U.S. government prior to the Tax Cuts and Jobs Act (TCJA) or as a minimum tax, such as the 15% agreement reached by the OECD, implementable by the end of 2023.

5. Make new discrete choice  $s_{t+1} \in \{E, D, X, M\}$ .
6. Realise iceberg cost state  $\tau_{t+1}^{s_{t+1}}$ . If upgrading to  $s_{t+1} \in \{X, M\}$  with  $s_t \neq s_{t+1}$ , assume  $\tau_H^{s_{t+1}}$ . If remaining in  $s_{t+1} \in \{X, M\}$  with  $s_t = s_{t+1}$ , then draw iceberg cost shock from process  $g_\tau(\tau_{t+1}^{s_{t+1}}, s_{t+1} | \tau_t^{s_t}, s_t)$ . If  $s_{t+1} = D$ , then assume  $\tau^D = \infty$ .

### 3.3.2 Entrants

Variables and parameters pertaining to new entrant firms will be denoted with  $T$  superscripts. These firms create a new variety upon startup. An entrant pays a sunk cost to incorporate their firm/establish their headquarters, which we denote by  $f^T$  (the associated cash outflow is  $W_t f^T$ ; see Section 3.5). Their initial status is as a domestic firm  $s_{t+1} = D$ . They draw their initial productivity shock from a distribution  $g^T(z)$  and take iceberg cost  $\tau_{t+1}^D = \infty$ .<sup>15</sup> They then operate as incumbent firms thereafter. We denote the mass of these new entrants by  $\Omega_t^T$ .

## 3.4 Government

The government acts passively. They raise taxes exogenously and re-distribute the proceeds to the household. The expression for aggregate tax collections is given as

$$T_t = (\hat{\tau}^X - 1)I_t + \hat{\tau}^M \Pi_t^M, \quad (12)$$

where the first term represents collected tariff revenues on the value of imports  $I_t$ , and the second is proceeds from taxing the  $H$  multinationals' profits  $\Pi_t^M$ .

## 3.5 Modelling of Fixed Costs

We now discuss the units of measure of the model's fixed costs. Trade models of firm heterogeneity typically denote fixed costs in terms of labour units. However there are complications with this in a model with imperfect labour markets. In particular, heterogeneous firms will all pay different wages. Moreover, new entrants must pay their sunk costs prior to drawing their productivity, as is standard in models of firm dynamics, so these costs cannot be denoted in terms of labour. One simple solution is to denote these fixed costs in terms of goods, as in [Jha and Rodriguez-Lopez \(2021\)](#) and previous versions of the present paper. However, this introduces additional inefficiencies into the environment. Specifically, it means that love of variety benefits that give gains to consumption will also lower the burden of fixed costs for firms; this yields a suboptimal firm mass. As mentioned above, our compromise is to denote these fixed costs in terms of final goods, but to index them to the aggregate wage index, with both the final good and wage index pertaining to the firm's headquarters country.<sup>16</sup> Below we rationalise this modelling choice in the context of a simplified static model. The static model also allows us to highlight the role of firm heterogeneity in driving the gains from trade in the monopsony model and the possibly adverse welfare implications of FDI.

<sup>15</sup>In the quantitative model, this distribution will be related to the productivity law of motion (11), to be discussed below.

<sup>16</sup>Note that with country symmetry, the results will be the same as indexing to the local wage index (except in the offshoring extension in Section 8.4).

## 4 Equilibrium of a Static Simplified Model

Before turning to the full model’s equilibrium behaviour, we first describe a simplified environment that permits analytical solutions. This theoretical model features a single time period; as such, we omit all time subscripts, while maintaining the interpretation of variables as defined previously. Policy instruments are set to zero — specifically,  $\hat{\tau}^X = 1$  and  $\hat{\tau}^M = 0$  — such that international engagement is governed solely by the physical iceberg cost  $\tau^s$  for  $s \in X, M$  (i.e., we assume  $\tau_L^s = \tau_H^s = \tau^s$ ). For open economy scenarios, we consider two symmetric countries; all liberalisation episodes are therefore bilateral. We consider the model’s predictions under perfect competition in labour markets ( $\theta \rightarrow \infty$ ), as well as under imperfect competition ( $\theta < \infty$ ).

**Roadmap.** The analysis and derivations in this section are intended to isolate the role of key model features. These serve to guide both the calibration strategy in Section 6 and the interpretation of quantitative results in Section 7.

1. *Fixed costs.* As discussed in Section 3.5, modelling fixed costs in terms of goods may introduce inefficiencies. To better understand this, we first consider a homogeneous-firm model with three alternative specifications for fixed costs, before proceeding with our preferred approach.
2. *Labour market power.* Firms exercising labour market power through wage markdowns ( $\theta/(1 + \theta)$ ) and the Frisch elasticity ( $\phi \geq 0$ ) introduce allocative inefficiencies.<sup>17</sup>
3. Role of firm heterogeneity.
4. Role of horizontal FDI.

### 4.1 Homogeneous-Firm Model

We begin by studying autarky in an economy with identical firms, under an arbitrary labour supply elasticity  $\theta$ . We construct a nested model that accommodates three alternative specifications for fixed costs: (i) fixed costs denominated in labour units ( $I = 1$ ), (ii) fixed costs specified in final-good units without wage indexation ( $I = 0, J = 0$ ), and (iii) fixed costs in final-good units indexed to the wage ( $I = 0, J = 1$ ). All firms offer a common wage  $w$ . The fixed continuation cost  $f^C$  is interpreted as follows: when  $I = 1$ , it is denominated in labour units; when ( $I = 0, J = 0$ ), it is denominated in units of the final good; and when ( $I = 0, J = 1$ ), it is equal to  $wf^C$  units of the final good.<sup>18</sup>

Given the production function for the final good in Equation (6) and the choice of numeraire,  $P = 1$ , the demand curve facing each intermediate-good producer is given by  $y = Ap^{-\sigma}$ . In the case where fixed costs are denominated in labour units ( $I = 1$ ), total firm labour demand can be written as  $n = \hat{n} + f^C$ , where  $\hat{n}$  denotes the variable labour input used in

---

<sup>17</sup>One can loosely gauge the effect of markdowns by comparing to an equilibrium where firms are subject to the upward-sloping labour supply curve (USLS), but fail to internalise it — i.e., firms behave myopically with respect to labour market conditions.

<sup>18</sup>We abstract from the sunk entry cost  $f^T$  in this section, since all firms draw the same productivity level. The expressions below can easily be modified to include  $f^T + f^C$ , with no loss of generality or insight.

production, and the remainder covers the fixed labour cost. This implies total output is given by  $y = z\hat{n} = z(n - f^C)$ . More generally, under the nested structure described above, firm revenue is given by  $A^{\frac{1}{\sigma}} z^{\frac{\sigma-1}{\sigma}} (n - If^C)^{\frac{\sigma-1}{\sigma}}$ . Firms maximise the following profit function, subject to the upward-sloping labour supply condition in Equation (3):

$$\max_n A^{\frac{1}{\sigma}} z^{\frac{\sigma-1}{\sigma}} (n - If^C)^{\frac{\sigma-1}{\sigma}} - wn - (1 - I)(Jw + (1 - J))f^C, \quad (13)$$

which yields the first-order condition

$$\frac{\sigma - 1}{\sigma} A^{\frac{1}{\sigma}} z^{\frac{\sigma-1}{\sigma}} (n - If^C)^{-\frac{1}{\sigma}} = \frac{1 + \theta}{\theta} w. \quad (14)$$

As usual, the first-order condition in (14) equates the marginal benefit and marginal cost of production. The former is shaped by the monopolistic markup  $\sigma/(\sigma - 1)$ , while the latter reflects the presence of a markdown  $\theta/(1 + \theta)$  arising from monopsonistic labour market power.

**Definition 1.** *An equilibrium is defined as an allocation in which all model agents optimise and the following conditions are satisfied:*

1. *Labour market clearing holds*

$$n = \left( \frac{\Gamma \bar{L}}{\Omega^P W^{\theta - \phi}} \right) w^\theta, \quad (15)$$

*where the left-hand side denotes firm-level labour demand, and the right-hand side corresponds to labour supply, in accordance with Equation (3).<sup>19</sup>*

2. *The CPI definition (8) holds.*

3. *Free entry of firms: the optimised value of equation (13) equals zero.*

In Appendix A.2, we prove the result stated below.

**Proposition 1.** *When fixed costs are denominated in units of labour and  $\phi \rightarrow 0$ , the decentralised equilibrium coincides with the planner's solution if firms do not exercise labour market power. The exercise of labour market power results in smaller firm employment and too many firms compared to the planner's solution.<sup>20</sup> By contrast, when fixed costs are specified in units of final output, the decentralised equilibrium is inefficient even when firms do not exercise labour market power and  $\phi \rightarrow 0$ .*

That is, when fixed costs are modelled in units of labour, the only distortion is the standard monopsony distortion arising from the exercise of labour market power. By contrast, when

<sup>19</sup>Under firm homogeneity, aggregate labour market clearing holds by simply scaling both sides by the number of active firms  $\Omega^P$ .

<sup>20</sup>When  $\phi > 0$ , the market outcome is distorted even when firms do not exercise labour market power because of the markup in the product market. That is, leisure is effectively subsidised relative to consumption, leading to under-supply of labour, as pointed out by Bilbiie, Ghironi and Melitz (2019).

fixed costs are specified in units of final output, there is a distortion even when firms do not exercise labour market power. This has the potential to quantitatively amplify misallocation through roundabout production, similarly to [Caliendo and Parro \(2015\)](#). Therefore, to stay closer to the modelling of fixed costs in labour units, we adopt a specification in which fixed costs are denominated in units of final output but indexed to the wage index in Equation (4). For example, all firms operating will pay  $wf^C$  from their cash flow to maintain their headquarters in the homogeneous firm case; this expression constitutes the corresponding demand in final goods market clearing. Indexing fixed costs in this way ensures they co-move with the average cost of labour in the economy, as captured by the wage index.

We now verify that, in the homogeneous firm setting, the gains from trade are identical under the two fixed cost specifications: (i) fixed costs in units of labour, and (ii) fixed costs in units of final output but indexed to the wage index. We then proceed to contrast these gains with those arising in the heterogeneous firm environment.

**Trade with homogeneous firms.** We characterise the trading equilibrium under the assumption that fixed costs are denominated in units of output but indexed to the wage index. A representative trading firm produces for both the domestic and foreign markets, denoted by  $q$  and  $q^*$ , respectively, and hires domestic labour  $n_D$  and  $n_X$  to produce goods sold in each market. The firm solves the following problem:

$$\max_{p,q,p^*,q^*,n_D,n_X} pq + p^*q^* - w(n_D + n_X),$$

subject to the demand curves in Equation (7), the labour supply curve in Equation (3), and the production function in Equation (10), where  $q = zn_D$  and  $\tau^X q^* = zn_X$ . Each firm also incurs fixed costs for both local and foreign operations, with total burden given by  $wf^X = w(f^C + f^{X,C})$ .

Using the equilibrium wage in autarky and under trade, we derive the implied gains from trade. We then compare these gains to the case where fixed costs are denominated in labour units and prove the following result in [Appendix A.3](#).

**Proposition 2.** *In the homogeneous firm model, the gains from trade are identical under both fixed cost specifications: when fixed costs are denominated in units of labour, and when they are specified in units of output but indexed to the wage.*

Denoting the own trade share by  $\lambda$ , we obtain  $\lambda = (1 + (\tau^X)^{1-\sigma})^{-1}$  and consequently, the trade elasticity is

$$-\frac{d \log \left( \frac{1-\lambda}{\lambda} \right)}{d \log \tau^X} = (\sigma - 1). \quad (16)$$

That is, the own trade share and trade elasticity are exactly the same as in a standard Krugman model with perfect labour market and do not depend on  $\theta$ . Additionally, the expression for the gains from trade written in terms of  $\lambda$ , derived in [Appendix A.4](#), is given by

$$\frac{w_X}{w_D} = \left( \frac{f^C}{f^X \lambda} \right)^{\frac{1}{\sigma-1-\phi}} > 1, \quad (17)$$

where  $w_X$  is the wage in the exporting equilibrium and  $w_D$  is the wage in autarky. That is, the gains from trade are independent of  $\theta$  as well.

The independence of the own trade share, the trade elasticity, and the gains from trade from  $\theta$  no longer holds in the heterogeneous-firm environment discussed below.

## 4.2 Heterogeneous-Firm Model

The key difference in the heterogeneous-firm case is that, due to productivity dispersion, firms pay different wages. As such, both fixed and sunk entry costs are indexed to the wage index  $W$ , defined in Equation (4). Upon paying the sunk cost  $Wf^T$ , firms draw their productivity  $z$  from a distribution with density  $g(z)$ .

In autarky, a firm's problem is analogous to that in the homogeneous case. However, the presence of fixed costs implies that only firms above a cutoff productivity level  $\widehat{z}_D$  are active. Firms with  $z < \widehat{z}_D$  cannot cover the fixed cost and exit. The cutoff  $\widehat{z}_D$  is determined by the free entry condition

$$\int_{\widehat{z}_D}^{\infty} \left[ \left( \frac{z}{\widehat{z}_D} \right)^{\chi} - 1 \right] f^C g(z) dz = f^T,$$

where the exponent  $\chi$  is a hyperparameter defined as  $\chi \equiv (\sigma - 1)(1 + \theta)/(\sigma + \theta)$ .

All derivations are provided in Appendix A.5, where we show that the wage index  $W$  — which serves as a sufficient statistic for welfare — is given by

$$W_A = \frac{(\sigma - 1)\theta}{\sigma(1 + \theta)} (\Omega_A^P)^{\frac{1}{\sigma-1}} \Phi_A, \quad (18)$$

where subscript  $A$  denotes “autarky”,  $\Omega_A^P$  is the mass of active firms, and  $\Phi_A$  is a weighted average of productivities across those firms, as defined in Appendix A.5.

**Trade with heterogeneous firms.** The profit maximisation problem of an exporting firm remains structurally similar to that under homogeneous firms. However, in the trading equilibrium two cutoffs arise: the survival cutoff  $\widehat{z}_D$  and the exporting cutoff  $\widehat{z}_X$ . These are pinned down by the following conditions:

$$\begin{aligned} \pi(\widehat{z}_D, D) &= Wf^C, \\ \pi(\widehat{z}_X, X) - Wf^C - Wf^{X,C} &= \pi(\widehat{z}_X, D) - Wf^C, \end{aligned}$$

where  $\pi(z, s)$  denotes the gross profit of a firm with productivity  $z$  and status  $s \in \{D, X\}$ .

The expression for welfare in the trading equilibrium is given by

$$W_T = \frac{(\sigma - 1)\theta}{\sigma(1 + \theta)} (\Omega_T^P)^{\frac{1}{\sigma-1}} \Phi_T, \quad (19)$$

where  $\Omega_T^P$  is the equilibrium mass of domestic firms, and  $\Phi_T$  is a weighted average of productivities across all firms selling in the country. Comparing Equations (18) and (19), the gains from trade in the heterogeneous-firm model are given by

$$\frac{W_T}{W_A} = \left( \frac{\Omega_T^P}{\Omega_A^P} \right)^{\frac{1}{\sigma-1}} \frac{\Phi_T}{\Phi_A}.$$

Assuming productivity follows a Pareto distribution with shape parameter  $k$ , the exporting cutoff can be written as  $\widehat{z}_X = \mu \widehat{z}_D$ , where

$$\mu \equiv \left[ \frac{\mathcal{F}}{(1 + (\tau^X)^{1-\sigma})^{\frac{\chi}{\sigma-1}} - 1} \right]^{\frac{1}{\chi}} \geq 1; \quad \mathcal{F} \equiv \frac{f^{X,C}}{f^C}.$$

The expression for the gains from trade becomes

$$\frac{W_T}{W_A} = \left( \frac{\mathcal{F} + \mu^k}{\mu^k} \right)^{\frac{1}{k} \frac{\sigma-1}{\sigma-1-\phi}} \left( \left( (1 - \mu^{\chi-k}) + (1 + (\tau^X)^{1-\sigma})^{\frac{1+\theta}{\sigma+\theta}} \mu^{\chi-k} \right)^{\frac{1}{\chi}} \right)^{\frac{\sigma-1-\chi}{\sigma-1-\phi}}$$

which is a non-linear function of  $\theta$ . We cannot analytically rank the welfare in autarky and trading equilibrium.

Turning to trade elasticities in the heterogeneous-firm model, the overall trade elasticity can be decomposed into intensive- and extensive-margin components. We prove the following result in Appendix A.6.

**Proposition 3.** *The intensive-margin trade elasticity increases with  $\theta$ , and — when productivity is Pareto distributed — the extensive-margin trade elasticity decreases with  $\theta$ .*

Since  $\theta \rightarrow \infty$  captures the benchmark Melitz model with perfect labour market, the proposition above implies that monopsony power in the labour market reduces the intensive-margin trade elasticity and increases the extensive-margin trade elasticity.

Several key insights emerge from the derivations in this section, particularly in contrast with the homogeneous-firm case discussed previously. Most importantly, firm heterogeneity is essential for monopsony power — governed by the parameter  $\theta$  — to influence the gains from trade. The presence of selection and a stronger extensive-margin response under lower  $\theta$  amplify the gains from trade due to love-of-variety effects in final-good aggregation.

This has two implications for the calibration strategy outlined in Section 6. First, the trade elasticity must be treated as a moment to target, to ensure a consistent basis for comparing the aggregate predictions of economies with and without labour market imperfections. Second, since the aggregate effects of  $\theta$  are mediated through the creation of new exporters — a margin activated by relatively few firms, and typically later in their lifecycle — reliable quantification requires targeting moments related to firm growth.

Before turning to the dynamic model and its empirical discipline, we conclude this static analysis by examining how FDI interacts with monopsony power in this simplified setting.

### 4.3 Exporting vs FDI in the Homogeneous-Firm Model

We now briefly study the case of FDI in the context of homogeneous firms. Suppose that foreign direct investment requires a fixed cost  $w f^{M,C}$  associated with operating a foreign plant. The total fixed cost for a multinational firm is then  $w f^M$ , where  $f^M = f^C + f^{M,C}$  reflects the combined costs of operating domestic and foreign plants.

In what follows, we abstract from FDI taxes ( $\widehat{\tau}^M = 0$ ) and assume no iceberg losses from foreign operations ( $\tau^M = 1$ ). Under these assumptions, firms produce identical quantities, face the same wage, and charge the same price in both markets.

Comparing wages — and thus welfare — across the two polar cases in which all firms either export ( $w_X$ ) or engage in FDI ( $w_M$ ), we obtain (derivation in Appendix A.7)

$$\frac{w_X}{w_M} = \left( \frac{f^M}{\underline{f} f^X} \right)^{\frac{1}{\sigma-1-\phi}}, \quad (20)$$

where  $\underline{f} \equiv \frac{2}{1+(\tau^X)^{1-\sigma}}$ . Note that the above wage ratio is independent of the labour supply elasticity,  $\theta$ . Equation (20) implies that welfare is higher under exporting relative to FDI when the following inequality holds:

$$\frac{f^M}{f^X} > \underline{f}. \quad (21)$$

In the absence of iceberg shipping costs, this reduces to  $w_X > w_M$  if and only if  $f^X < f^M$ . However, when shipping costs are positive, welfare may still be higher under FDI even if  $f^X < f^M$ .

Importantly, in the presence of monopsonistic labour markets, firms have an incentive to engage in FDI in order to suppress wages. As a result, an equilibrium featuring FDI may emerge even when the inequality in (21) is satisfied — that is, even when welfare would be higher under exporting.

In Appendix A.8, we show that there exists a cutoff value  $\bar{f}$  given by

$$\bar{f} \equiv \frac{2}{[1 + (\tau^X)^{1-\sigma}]^{\frac{1+\theta}{\sigma+\theta}}},$$

such that an FDI equilibrium arises if  $\frac{f^M}{f^X} < \bar{f}$ , and an exporting equilibrium otherwise. We also verify that  $\bar{f} > \underline{f}$  for any finite  $\theta$ . Thus, there exists a region of the parameter space in which an FDI equilibrium emerges despite delivering lower welfare than the exporting alternative. We summarise this result in the proposition below.

**Proposition 4.** *If the ratio of FDI to export fixed costs satisfies*

$$\underline{f} < \frac{f^M}{f^X} < \bar{f},$$

*then the equilibrium features all firms engaging in FDI. However, welfare is lower than in a counterfactual scenario in which firms are restricted to exporting.*

Facing upward-sloping labor supplies in each market, the reason firms may prefer FDI over exporting is to suppress wages. However, this incentive to escape their local production structure does not necessarily align with social welfare.

**Taking stock.** Before turning to the equilibrium of the full dynamic model, we briefly summarise the key insights from the static framework, in one-to-one correspondence with the roadmap presented at the start of this section.

1. *Fixed costs.* The main inefficiency in the model is the monopsony distortion, but the roundaboutness in production created by modelling fixed cost in terms of final output amplifies that inefficiency. We neutralise this additional roundaboutness-driven wedge by indexing fixed costs to the average wage.
2. *Wage markdowns and Frisch elasticity  $\phi \geq 0$  as potential inefficiencies.* We retain firm wage-setting power and a positive Frisch elasticity in our baseline specification to maintain empirical relevance. Their quantitative impact is evaluated in Section 8.
3. *Role of firm heterogeneity.* In the homogeneous-firm setting, the welfare response to trade liberalisation is identical across labour market structures. With heterogeneity, however, imperfect labour markets generate a stronger extensive-margin response, amplifying gains from trade.
4. *Role of horizontal FDI.* Monopsony creates an inefficiency wherein firms undertake FDI to avoid rising marginal labour costs associated with exporting, thereby suppressing wages. As this margin is only accessed by a small subset of firms, and typically later in their lifecycle, our calibration strategy must target firm growth moments conditional on status  $s \in \{D, X, M\}$ .

## 5 Equilibrium of the Dynamic Quantitative Model

We now return to the full dynamic model specified in Section 3, towards the goal of quantifying the simple insights of Section 4 and how they vary over time.

### 5.1 Intermediate Goods Firms

We now present the recursive formulations of the incumbent and entrant's problems. We leave time subscripts on functions and variables given the transition analysis in the quantitative results.

#### 5.1.1 Incumbents

The value function for an incumbent firm at the time of production — Step 3 of the timing part of Section 3.3.1 — with state  $(z_t, s_t, \tau_t^{s_t})$  is given by

$$v_t(z_t, s_t, \tau_t^{s_t}) = \pi_t(z_t, s_t, \tau_t^{s_t}) + \mathbb{E}_t \max_{s_{t+1} \in \{E, D, X, M\}} \tilde{v}_t(z_t, s_t, \tau_t^{s_t}, f_t, s_{t+1}) \quad (22)$$

where  $\pi_t(z_t, s_t, \tau_t^{s_t})$  is the period profits given current status  $s_t \in \{D, X, M\}$ ; these all depend on the firm's  $z_t$  and  $\tau_t^{s_t}$  draws. The term  $\tilde{v}_t(z_t, s_t, \tau_t^{s_t}, f_t, s_{t+1})$  is the continuation value of the firm, conditional on choosing a status  $s_{t+1}$ . Notice that this function depends on the draws of the firms' sunk costs, captured by vector  $f_t$  and accounted for with the expectation operator. We denote the policy function for the optimal status choice by  $s_{t+1}(z_t, s_t, \tau_t^{s_t}, f_t)$ . The conditional value from exiting is given simply by  $\tilde{v}_t(z_t, s_t, \tau_t^{s_t}, f_t, E) = 0$ ; the firm ceases to exist. A non-exit choice ( $s_{t+1} \in \{D, X, M\}$ ) yields conditional value

$$\tilde{v}_t(z_t, s_t, \tau_t^{s_t}, f_t, s_{t+1}) = -W_t f_t(s_{t+1}, s_t) + \beta \mathbb{E}_t[v_{t+1}(z_{t+1}, s_{t+1}, \tau_{t+1}^{s_{t+1}})], \quad (23)$$

where  $f_t(s_{t+1}, s_t)$  represents the entry in vector  $f_t$  for a firm moving from status  $s_t \in \{D, X, M\}$  to  $s_{t+1}$  (as per previous discussion, it is indexed to wages). Equation (23) says that the firm pays the corresponding fixed cost combination for their choice, based on their current status and drawn sunk cost shocks, then receives the corresponding expected discounted continuation value. The expression for period profits is then given by

$$\begin{aligned} \pi_t(z_t, s_t, \tau_t^{s_t}) = & \max_{\{k_t(z_t, s_t, \tau_t^{s_t})\}} p_t(z_t, s_t, \tau_t^{s_t})q_t(z_t, s_t, \tau_t^{s_t}) + \mathbb{1}_{s_t=X} [p_t^*(z_t, X, \tau_t^X)q_t^*(z_t, X, \tau_t^X)] \\ & - w_t(z_t, s_t, \tau_t^{s_t})n_t(z_t, s_t, \tau_t^{s_t}) \\ & + \mathbb{1}_{s_t=M}(1 - \hat{\tau}^M) [p_t^*(z_t, M, \tau_t^M)q_t^*(z_t, M, \tau_t^M) - w_t^*(z_t, M, \tau_t^M)n_t^*(z_t, M, \tau_t^M)], \end{aligned} \quad (24)$$

where  $k_t(z_t, s_t, \tau_t^{s_t})$  is an array of static control variables, which depend on the firm's state and  $\mathbb{1}_x$  is an indicator function that equals 1 when  $x$  is true and 0 otherwise. Note that the maximisation in (24) is done subject to firms' technology (10), the household's labour supply (3) and the final goods demand (7), where the number of constraints varies based on the firm's status.

The first expression on the right-side of equation (24) is the sales the firm generates domestically. The second term represents the extra sales income it receives, should it choose to export some of its output. The third term is the wage bill the firm incurs domestically. Note that, holding  $z_t$  constant, the first and third terms may differ between a domestic firm and exporting firm due to the upward-sloping labour supply curve. The fourth term in (24) is the profits from overseas operations, should the firm have multinational status, upon which they will pay the outward FDI tax  $\hat{\tau}^M$ . The array  $k_t(z_t, s_t, \tau_t^{s_t})$  consists of  $p_t(z_t, s_t, \tau_t^{s_t}), q_t(z_t, s_t, \tau_t^{s_t}), w_t(z_t, s_t, \tau_t^{s_t}), n_t(z_t, s_t, \tau_t^{s_t})$  for all three statuses, with the addition of  $p_t^*(z_t, s_t, \tau_t^{s_t}), q_t^*(z_t, s_t, \tau_t^{s_t})$  for  $s_t \in \{X, M\}$  and also  $w_t^*(z_t, s_t, \tau_t^{s_t}), n_t^*(z_t, s_t, \tau_t^{s_t})$  for  $s_t = M$ . Lastly, note that the idiosyncratic iceberg cost shock  $\tau_t^s$  features directly in the expressions for these optimal controls for status  $s$ , affecting the labour input requirement and hence the choices of prices and output.

### 5.1.2 Entrants

An entrant receives value function  $v_t^T \geq 0$  of the form

$$v_t^T = -W_t f^T + \beta \mathbb{E}_t^T [v_{t+1}(z_{t+1}, D, \tau^D = \infty)], \quad (25)$$

where notice that the expectation is with respect to the entrants' draw from the (potentially truncated) unconditional distribution of productivity. The entrant incurs the fixed cost at time  $t$  and then receives the discounted expected value from being an incumbent with domestic status. This value must be weakly positive for a positive mass of entrants to exist; this mass is zero otherwise.

## 5.2 Cross-Sectional Measure

We track the cross-section of firms headquartered in each country along the state space  $(z_t, s_t, \tau_t^{s_t})$ . Denote the corresponding cross-section of firms in period  $t$ , before the sunk cost

shocks are drawn, by  $\mu_t(z_t, s_t, \tau_t^{s_t})$ . The law of motion for this measure is given as follows

$$\begin{aligned} \mu_{t+1}(z_{t+1}, s_{t+1}, \tau_{t+1}^{s_{t+1}}) &= \Omega_t^T \mathbb{1}_{s_{t+1}=D} \mathbb{1}_{\tau_{t+1}^{s_{t+1}}=\tau^D=\infty} g^T(z_{t+1}) \\ &+ \sum_{s_t \in \{D, X, M\}} \sum_{\tau_t^{s_t} \in \{\tau_H^{s_t}, \tau_L^{s_t}\}} \int_{z_t} \int_{f_t} \mathbb{1}_{s_{t+1}=s_{t+1}(z_t, s_t, \tau_t^{s_t}, f_t)} g(z_{t+1}|z_t) g_\tau(\tau_{t+1}^{s_{t+1}}, s_{t+1}|\tau_t^{s_t}, s_t) H(f_t) \\ &\quad \times \mu_t(z_t, s_t, \tau_t^{s_t}) dz_t df_t \end{aligned} \quad (26)$$

for any arbitrary  $z_t, z_{t+1}, \tau_t^{s_t}, \tau_{t+1}^{s_{t+1}}$  and  $s_t, s_{t+1} \in \{D, X, M\}$ . The first term in equation (26) reflects that entrants — of mass  $\Omega_t^T$  — enter at status  $s_{t+1} = D$  by design and draw their initial productivity shock from unconditional distribution  $g^T(z_{t+1})$ . The second term says that the incumbent firms draw from the sunk cost probability process  $H(f_t)$ , upon which their status choice  $s_{t+1}(z_t, s_t, \tau_t^{s_t}, f_t)$  is conditional. They then draw their iceberg cost shock  $\tau_{t+1}^{s_{t+1}}$  in accordance with their choice, before drawing productivity  $z_{t+1}$  at the start of  $t + 1$ . We then integrate and sum over all possible  $z_t, f_t, s_t$  and  $\tau_t^{s_t}$  values that can lead to arbitrary state  $(z_{t+1}, s_{t+1}, \tau_{t+1}^{s_{t+1}})$  at  $t + 1$  to obtain  $\mu_{t+1}(z_{t+1}, s_{t+1}, \tau_{t+1}^{s_{t+1}})$ . We can then write the overall measure of locally-headquartered firms as

$$\begin{aligned} \Omega_t &= \sum_{s_t \in \{D, X, M\}} \sum_{\tau_t^{s_t} \in \{\tau_H^{s_t}, \tau_L^{s_t}\}} \int_{z_t} \mu_t(z_t, s_t, \tau_t^{s_t}) dz_t \\ &= \Omega_t^D + \Omega_t^X + \Omega_t^M, \end{aligned} \quad (27)$$

meaning that the total mass of locally-headquartered firms at the time of production is the set of domestic ( $\Omega_t^D$ ), exporting ( $\Omega_t^X$ ), and multinationals ( $\Omega_t^M$ ) — status chosen from last period — integrated over all other idiosyncratic states.<sup>21</sup> Note also that the exiting firm mass  $\Omega_t^E$  can be derived from the difference of  $\Omega_t$  minus the second term (the sums/integrals) on the right-side of equation (26).

### 5.3 Equilibrium Definition

We define equilibrium as an infinite sequence of equilibrium objects for country  $H$ . Specifically, a list of cross-sectional objects

$$\left\{ \left\{ v_t(z_t, s_t, \tau_t^{s_t}), \mu_t(z_t, s_t, \tau_t^{s_t}), k_t(z_t, s_t, \tau_t^{s_t}) \right\}_{(z_t, s_t, \tau_t^{s_t})}, \left\{ s_{t+1}(z_t, s_t, \tau_t^{s_t}, f_t) \right\}_{(z_t, s_t, \tau_t^{s_t}, f_t)} \right\}_{t=0}^\infty,$$

as well as aggregates

$$\{A_t, P_t, W_t, N_t, \Pi_t, C_t, F_t, T_t\}_{t=0}^\infty, \quad (28)$$

the collection of which we denote as  $\Upsilon$ . Infinite sequence  $\Upsilon$  is such that the following conditions are satisfied:

1. Workers are optimising.
2. The value function  $v_t(z_t, s_t, \tau_t^{s_t})$  and the set of policy functions for control variables  $k_t(z_t, s_t, \tau_t^{s_t}), s_{t+1}(z_t, s_t, \tau_t^{s_t}, f_t)$  solve firm problem (22) with all aggregate objects taken as given.

<sup>21</sup>Where mass status  $s \in \{D, X, M\}$  is defined as  $\Omega_t^s = \sum_{\tau_t^s \in \{\tau_H^s, \tau_L^s\}} \int_{z_t} \mu_t(z_t, s, \tau_t^s) dz_t$ .

3. Cross-sectional measure  $\mu_t(z_t, s_t, \tau_t^{st})$  satisfies the law of motion given in equation (26).
4. The free entry condition holds:  $v_t^T = 0$  where value to entry is given by equation (25).
5. Aggregate objects can be found using the cross-sectional measure  $\mu_t(z_t, s_t, \tau_t^{st})$ . Market clearing in the final goods market gives

$$A_t = C_t + W_t F_t,$$

where  $A_t$  is given by equation (6),  $C_t$  is given by the household objective in (5), and  $F_t$  is aggregate fixed costs used by  $H$ -headquartered firms. Again notice that these fixed costs are indexed to the wage index  $W_t$ . Aggregate object  $\Pi_t$  is aggregate profits and  $T_t$  is aggregate tax collections.

6. The government budget constraint holds as in equation (12), giving  $T_t$ .
7. Labour market clearing holds, both at the individual firm level and in the aggregate.

The conditions for  $F$  are defined similarly with infinite sequence denoted by  $\Upsilon^*$ ; the conditions in  $H$  and  $F$  all hold simultaneously in any equilibrium. Note that the conditions in Definition 1, corresponding to the simplified static model, are nested within the structure of the dynamic equilibrium defined above.

## 6 Calibration

In this section, we detail the parameter choices used in the numerical solution. We consider several variations of the model and parameter space to quantify the impact of labour market monopsony on openness reforms. We begin by examining a tariff reduction in calibrations where FDI is prohibitively costly — that is,  $\widehat{f}^{M,D} = \widehat{f}^{M,X} = f^{M,C} = \infty$  — in order to isolate the effect of imperfect labour markets on the domestic–foreign engagement margin. We then extend the analysis to incorporate the export–FDI substitution margin in the context of multinational tax reforms. For the purpose of studying trade liberalisations, we start with the following two sets of calibrated parameters, where  $C$  stands for calibration.

- $C_I$ : the baseline imperfect labour markets structure. Firms are wage-setters and face USLS with  $\theta < \infty$ .
- $C_P$ : the perfect labour markets structure. Since  $\theta = \infty$ , firms hire as much labour as they like at a fixed, uniform market wage.

Note that each time we change the structure of the model across  $C_I$  and  $C_P$ , we re-calibrate so that the behaviour of firms in the model matches a common set of moments in the data.<sup>22</sup> Comparing the counterfactual results of  $C_I$  to  $C_P$  gives monopsony’s quantitative impact. We denote the equivalent calibrations in the FDI parameterisation as  $C_I^{FDI}$  and  $C_P^{FDI}$ . Note also that the two countries are taken to be symmetric across all calibrations.

---

<sup>22</sup>Given that our focus is on quantification of the labour market frictions’ effect on a policy *change*, it is necessary to hold the starting point constant. By re-calibrating, we ensure that firms’ behaviour in the pre-reform economy is always the same. Some examples of the many other papers taking this approach are Costinot and Rodríguez-Clare (2014), Atkeson and Burstein (2010), Simonovska and Waugh (2014) and Brooks and DAVIS (2020). For a more qualitative discussion on how comparative statics on  $\theta$  affect the gains from trade, absent re-calibration, see Jha and Rodríguez-Lopez (2021).

Parameter	Symbol	Value	Source
Discount factor	$\beta$	0.98	Literature
Frisch elasticity	$\phi$	0.20	Literature
Elasticity of labour supply	$\theta$	1.88	Literature
Elasticity of substitution	$\sigma$	5.00	Literature
Tariff with imperfect markets ( $C_I$ )	$\hat{\tau}^X$	1.10	Literature
FDI tax with imperfect markets ( $C_I^{FDI}$ )	$\hat{\tau}^M$	0.01	Normalisation
Persistence of productivity	$\rho_z$	0.66	Compustat
Variability of productivity	$\sigma_z$	0.22	Compustat

Table 1: Parameters calibrated outside the model.

To keep the computational burden low and to ensure proper identification, we have two sets of parameters. One set are for parameters chosen outside the model, directly from the data and previous studies in the literature. The second set consists of parameters calibrated inside the model to ensure consistency of firm-level model moments with the data.

Table 1 shows the parameter values selected outside the model as well as their source. The discount factor is taken to be consistent with real interest rates in recent years. This is also the value used in several papers studying firm dynamics (e.g. see [Corbae and D’Erasmus, 2021](#)). We take the Frisch elasticity as the lower bound of those considered in [Berger, Herkenhoff and Mongey \(2022\)](#).<sup>23</sup> We take the elasticity of the firm-level labour supply to be 1.88, the mean firm-level empirical estimate from [Yeh, Macaluso and Hershbein \(2022\)](#).<sup>24</sup> We take the elasticity of substitution across varieties in final-good aggregation as 5 as in [Alessandria, Choi and Ruhl \(2021b\)](#). The parameters for the firm productivity process are estimated from firm-level data in Compustat over the period 1979–2018. We use the estimation method in [Cooper and Haltiwanger \(2006\)](#) to this end, yielding estimates of 0.66 for the persistence parameter and 0.22 for the variability parameter. We then discretise the productivity process into a Markov process using the method of [Tauchen \(1986\)](#). We then calibrate the remaining parameters inside the model using simulated method of moments, such that model moments align quantitatively with the data.

We give some detailed discussion now on two parameters that are crucial to the quantification exercises — the policy instrument magnitudes  $\hat{\tau}^X$  and  $\hat{\tau}^M$ . Recall from discussion of Proposition 3 in the simple model that the trade elasticity can potentially be varying, in the face of firm heterogeneity, across parameterisations with differing values of  $\theta$ . Guided by this, we calibrate the tariff rate inside the model in order to hold the long-run trade elasticity constant in the tariff liberalisation counterfactuals. Specifically, as in [Alessandria, Choi and Ruhl \(2021b\)](#), we set this parameter to be 10% in the  $C_I$  calibration. We then treat the long-run trade elasticity that follows in  $C_I$  (when we run the counterfactual in Section 7.1) as a moment to be matched in the calibration of  $C_P$ , using the tariff parameter as the

<sup>23</sup>The authors find that the welfare losses of labour market power are increasing in the Frisch elasticity. As such, this value will give more conservative estimates of monopsony’s effect on openness reforms, while still retaining empirical relevance.

<sup>24</sup>There is a large spread in empirical estimates of this elasticity; this value sits within this range. For a comprehensive discussion of the empirical literature, see [Jha and Rodriguez-Lopez \(2021\)](#).

corresponding instrument.<sup>25</sup> More details on the elasticity are provided in Table 4.

As an alternative to holding the trade elasticity constant, one could instead fix the steady-state ratio of tariff revenue collections to final output constant across calibrations using this parameter. We did this in previous versions of the paper and have verified with the current parameterisations that it yields effectively the same quantitative results; we then use this as justification for the calibration of  $\hat{\tau}^M$ . Specifically by normalising the FDI tax rate to 1% in  $C_I^{FDI}$ , and then treating the ratio of FDI taxes collected to final output as a moment to be matched in  $C_P^{FDI}$ .

We then calibrate the remaining 15 parameters internally so that the model moments match data analogues. Table 2 shows the value of each parameter under each calibration — without FDI,  $C_I$  and  $C_P$ , and with FDI,  $C_I^{FDI}$  and  $C_P^{FDI}$  — as well as the target moment that identifies it. Table 3 shows the match of the model to the data moments.

We now briefly turn to describe the identification of the parameters. It is important to note that, since this is a general equilibrium model, there is not a one-to-one mapping of any given parameter to a data moment. Instead, a change in one parameter will likely impact several moments. We therefore choose target moments that will shed the most light on the parameter of interest. Since it is difficult to disentangle parameters  $f^C$  and  $f^T$  in the data, we follow standard practice in the firm dynamics literature and set one of the two to normalise an aggregate variable to unity (e.g. the wage in [Hopenhayn and Rogerson, 1993](#)). In particular, we set the sunk entry cost ( $f^T$ ) to normalise the pre-liberalisation supply shifter in the labour supply curve —  $B$  in Equation (3) — to one.

We draw upon the summary statistics regarding foreign engagement statuses of establishments from [Boehm, Flaaen and Pandalai-Nayar \(2020\)](#). This study works with U.S. census data and summarises key moments relating to employment and establishment counts. They classify observations as pertaining to domestic, exporting, importing, U.S. multinational, and foreign multinational establishments.<sup>26</sup> The specific moments we borrow are transition probabilities across statuses, to identify the mean of the corresponding sunk cost draws. We then take the fractions of exporters and multinationals to identify the associated period-by-period fixed costs. A higher variability in the sunk cost distribution gives larger spread in the productivity of firms upgrading to  $X$  or  $M$  status. As such, we identify this parameter by matching the exit rate of exporting firms. The fixed continuation cost for all firms to maintain their headquarters,  $f^C$ , targets the economy-wide exit rate from the Business Dynamics Statistics (BDS) dataset.

---

<sup>25</sup>The long run elasticity numbers we find are similar to those from [Alessandria, Choi and Ruhl \(2021b\)](#). When setting the tariff rate to 10% in  $C_P$  without targeting the elasticity, we find a long-run elasticity of around 9. Our calibration procedure therefore must bring this number up to around 12/13 as in  $C_I$ .

<sup>26</sup>One issue with using these data is that ours is a model of firms, rather than establishments. We treat these establishment moments as an approximation for the transitions experienced by firms across statuses. Previous versions of this paper have instead calibrated directly to firm-level transitions using Compustat segment information and the procedure of [Fillat and Garetto \(2015\)](#) (or alternatively classifying multinational firms based on their reporting of pretax foreign income, e.g. see [Spencer, 2024](#)). Since Compustat pertains to large, publicly-listed firms, transitions are much rarer than in the population of firms at large, while also over-representing multinationals in the stationary equilibrium (compared with [Bernard, Jensen and Schott, 2009](#)). Aggregate insights from comparing across the different calibrations are largely robust to which approach is used.

		Trade		FDI		
		Imperfect	Perfect	Imperfect	Perfect	Target
Sunk entry ( $T$ ) cost	$f^T$	0.35	0.47	0.30	0.45	Unit supply-shifter $B$
$T$ productivity truncation	$\bar{z}^T$	0.89	1.00	0.81	1.00	3 year ave $T$ growth perfect markets
Fixed cost headquarters	$f^C$	0.23	0.16	0.24	0.17	Exit rate
Sunk cost ( $D, X$ )	$\hat{f}^{X,D}$	0.02	0.20	0.04	0.08	Transition ( $D, X$ )
Fixed cost $X$	$f^{X,C}$	0.02	0.03	0.02	0.04	Fraction $X$
Sunk cost std dev	$\sigma_f$	0.14	0.21	0.21	0.26	Transition ( $X, E$ )
$X$ iceberg cost 1	$\tau_H^X$	1.71	1.82	1.74	1.76	1 year export intensity
$X$ iceberg cost 2	$\tau_L^X$	1.15	1.18	1.10	1.14	Overall export intensity
$X$ iceberg cost persistence	$\rho_\tau^X$	0.85	0.53	0.87	0.71	3 year ave export growth
Tariff	$\hat{\tau}^X$	1.10	1.08	1.10	1.10	Trade elasticity imperfect markets
Sunk cost ( $X, M$ )	$\hat{f}^{M,X}$			0.09	0.09	Transition ( $X, M$ )
Sunk cost ( $M, X$ )	$\hat{f}^{X,M}$			0.03	0.04	Transition ( $M, X$ )
Fixed cost $M$	$f^{M,C}$			0.13	0.12	Fraction $M$
$M$ iceberg cost 1	$\tau_H^M$			2.28	1.48	1 year FDI intensity
$M$ iceberg cost 2	$\tau_L^M$			1.19	1.02	Overall FDI intensity
$M$ iceberg cost persistence	$\rho_\tau^M$			0.81	0.61	3 year ave FDI growth
FDI tax	$\hat{\tau}^M$			0.01	0.02	Taxes/output imperfect markets

Table 2: Parameters calibrated inside the model to match moments. Fixed costs are expressed as corresponding expenditure divided by total output.

	Trade			FDI		
	Target	Imperfect	Perfect	Target	Imperfect	Perfect
Exit rate	8.00	7.79	7.60	8.00	8.47	10.2
Transition ( $D, X$ )	5.40	4.93	3.27	5.40	3.83	6.26
Transition ( $X, E$ )	5.40	3.65	5.79	5.30	4.73	5.55
Fraction $X$	36.0	36.8	36.1	29.0	28.0	27.8
Export intensity	15.0	15.7	16.1	15.0	16.7	16.1
1 year export intensity	7.50	6.72	5.91	7.50	6.41	6.13
3 year ave export growth	17.5	17.4	17.4	17.5	17.5	17.5
3 year ave entrant growth	5.55	5.55	5.57	8.25	8.25	8.31
Trade elasticity	12.98	12.98	12.30			
Transition ( $X, M$ )				0.80	3.15	4.08
Transition ( $M, X$ )				1.90	0.98	2.34
Fraction $M$				5.50	4.58	5.22
FDI intensity				30.0	30.8	32.1
1 year FDI intensity				20.0	20.2	17.4
3 year ave FDI growth				7.50	7.57	7.27
FDI taxes/output				0.04	0.04	0.05

Table 3: Calibration for model variants. All numbers are after multiplication by 100 except for the trade elasticity.

We match several moments relating to the lifecycle of firm size. For the lifecycle of exporters, we calibrate the two stochastic iceberg cost values  $\tau_H^X$  and  $\tau_L^X$  to match moments related to the distribution of export intensities. We match an average export intensity of 15%, across the entire distribution of such firms. Then we borrow a data fact from [Alessandria, Choi and Ruhl \(2021b\)](#) relating to the intensity of first-year exporters — namely that it sits at half of that for the population at large. For the persistence of the iceberg cost process,  $\rho_\tau^X$ , we target the growth rate in the level of foreign income over the first three years of export operations, taken from [Gumpert, Li, Moxnes, Ramondo and Tintelnot \(2020\)](#). We follow a similar approach for calibrating the multinational lifecycle. For the two physical iceberg costs  $\tau_L^M$  and  $\tau_H^M$ , we match the average foreign sales intensity in the data of 30%, as well as an approximated value for the intensity of a first year multinational.<sup>27</sup> Then for  $\rho_\tau^M$ , we take the average growth rate in the level of FDI income over a firm’s first three years as a multinational ([Gumpert, Li, Moxnes, Ramondo and Tintelnot, 2020](#)). Finally, we adjust the truncation of the entrants’ productivity draws,  $\bar{z}^T$ , such that the growth rate of startups is held constant across calibrations.

<sup>27</sup>[Garetto, Oldenski and Ramondo \(2019\)](#) state that new multinationals typically jump straight to their long-run sales intensity, with respect to *horizontal* sales. This intensity then typically only rises, in a discrete fashion, when the subsidiary commences exporting as a platform. Once it starts exporting, the intensity typically rises by around 1/3, hence we take the starting intensity to be 2/3 of that in the long-run.

## 7 Quantitative Exercises

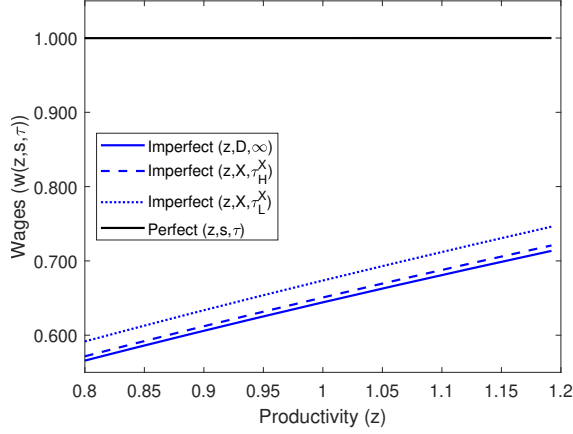
### 7.1 Export Dynamics and Trade Liberalisation

How does monopsony affect the lifecycle of exporters and firms more generally, and what are the resulting implications for tariff liberalisation? First, we examine the implications of this restriction for the pre-reform steady state. With this intuition in place, we then consider episodes of tariff liberalisation. In the second step, we explore how monopsony shapes changes in cross-sectional moments — such as firm survival and growth — in response to a tariff cut. Finally, in the third step, we examine how these cross-sectional effects influence the transition dynamics of aggregate variables in the counterfactual experiments.

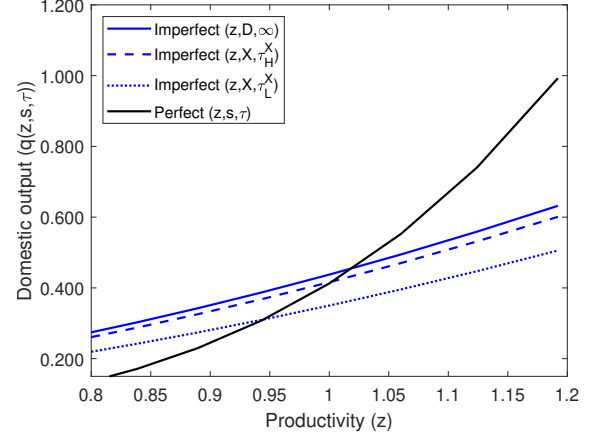
**Pre-reform steady state.** How does monopsony power affect a firm’s decision to begin exporting? Exporting typically occurs later in a firm’s lifecycle — only around 5% transition from domestic to exporting status in any given period, both in the data and the calibration. It is thus natural to ask how a firm’s observed decisions change in the period when it first begins serving a foreign market, as well as in the years that follow. Figure 1 depicts several sets of policy functions for various stages of the lifecycle — for wages ( $w(z, s, \tau)$ , 1a), domestic output ( $q(z, s, \tau)$ , 1b) and foreign output ( $q^*(z, s, \tau)$ , 1c). In these three policy function figures, solid lines correspond to domestic ( $D$ ) firms, dashed lines to young exporters ( $X$  with  $\tau_H^X$ ), and dotted lines to mature exporters ( $X$  with  $\tau_L^X$ ). Figures 1a–1c depict choices as a function of productivity, for all combinations of status and iceberg cost draws.

Figure 1a shows that monopsony leads to a wage lifecycle, which varies with both productivity and firm status. This variation is absent in  $C_P$ , where all firms pay the market wage rate of unity. A firm’s productivity  $z$  moves around over its lifetime: this can be thought of as movements left or right on the horizontal axes. As  $z$  increases, the firm optimally expands in size, meaning it must offer higher wages in  $C_I$  to attract more workers. In contrast with the  $z$  process, which gives continuous movements in the wage rate, upgrading from  $s = D$  to  $s = X$  causes a discrete shift in the wage policy function. As a firm becomes a young exporter (from the solid to the dashed line), it must offer higher wages at all levels of productivity. With this decision, the firm hires labour for production destined for domestic and foreign sales from the same upward-sloping labour supply schedule. When the firm eventually becomes a mature exporter (moving from the dashed to dotted line), the marginal product of its employees in export production rises, giving a further expansion in size, placing further upward-pressure on wages across the board. The magnitude of this latter shift is relatively strong given the large difference in the associated iceberg costs —  $\tau_H^X = 1.71$  and  $\tau_L^X = 1.15$ , respectively.

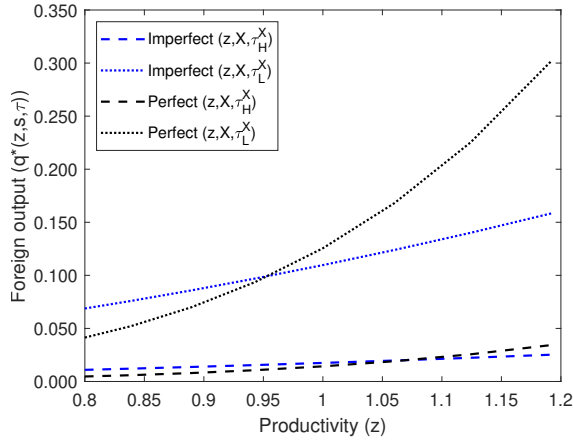
This wage lifecycle under monopsony has implications for marginal costs and for output sold locally (in  $H$ ) and abroad (to  $F$ ), as shown in Figures 1b and 1c, respectively. There are two main points to notice when comparing these policy functions for  $C_I$  and  $C_P$ . First, the slopes of the policy functions in each market are greater in  $C_P$  than for  $C_I$ ; under the latter, increasing marginal cost with monopsony gives stronger persistence of a firm’s production over time. In contrast, the constant marginal cost in  $C_P$  allows for more drastic adjustments to the intensive margin in response to productivity shocks. Second, upgrading from  $s = D$



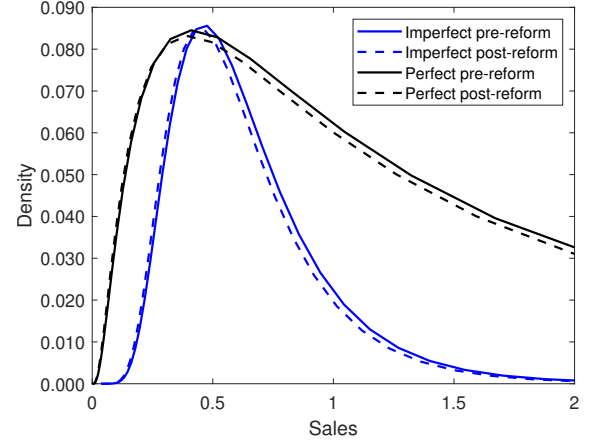
(a) Steady-state wage policy functions



(b) Steady-state  $H$  output policy functions



(c) Steady-state  $F$  output policy functions



(d) Firm size distribution

Figure 1: Steady-state lifecycle decision rules and firm distribution. Note that the sales distribution is presented as a fraction of final output. Note that these figures are for only a subset of the range of possible values for productivity  $z$  and sales.

to  $s = X$  in  $C_I$  causes production for  $H$  to decline — an effect that intensifies as the export branch matures. By contrast, in  $C_P$  the output policy function for  $H$  remains invariant across firm statuses. Given the single upward-sloping labour supply curve, the cost structures for production for  $H$  and  $F$  sales are non-separable with monopsony — an expansion in supply to  $F$  will raise the marginal cost of  $H$  production.

Lastly, Figure 1d shows pre-reform differences in the firm-size distribution across  $C_I$  and  $C_P$ , measured by sales. As a feature that follows from the discussion of output policy functions, the distribution of sales is considerably compressed in the presence of monopsony. Increasing marginal cost acts as a constraint on the expansion of the most productive firms — in  $C_I$ , the firm at the 67<sup>th</sup> percentile has sales that are 50% higher than the firm in the 33<sup>rd</sup> percentile. This contrasts with  $C_P$ , where constant marginal cost allows for productive firms to scale-up considerably, where the former firm is 150% larger than the latter. Together, these pre-reform

steady-state insights point to monopsonistic labour markets being associated with a lower elasticity of firm size to profitability. Quantitatively, the elasticity of a firm’s employment to a productivity  $z$  change, holding fixed its status and iceberg cost, is over four times larger in  $C_P$  than  $C_I$ . This difference remains roughly the same, should the productivity improvement prompt the firm to upgrade from  $D$  to  $X$  status. As such, this gives suggestive evidence that the magnitude of any *individual firm’s response* to a trade policy change will likely be muted in  $C_I$  relative to  $C_P$ .

**Liberalisation episodes and the firm cross-section.** Analysis of the steady state pointed towards the idea that monopsony dampens firm-level responses to profitability shocks. What does this imply for how the growth and survival trajectories of new firms and exporters respond to trade liberalisation?<sup>28</sup> To that end, we implement a bilateral tariff reduction of its calibrated value to zero in each of the two parameterisations,  $C_I$  and  $C_P$ .

Figure 2 depicts firm lifecycle trajectories, showing percentage point (pp) differences in key variables related to firm expansion between the post-liberalisation and pre-liberalisation steady states. For instance, Figure 2c shows the effect of liberalisation on the expected sales growth rate of new entrant firms across the different calibrations. In constructing this figure, we solve for the sales growth rate from one year to the next at each of the first 15 years of a firm’s life, in both the pre- and post-liberalisation steady states. The number shown for each year gives the pp difference between the growth rate in the post-reform steady state and that in the pre-reform steady state. For instance,  $C_I$  predicts that an entrant sales growth after its first year of operation will be 2pp higher with the tariff reduction. Given that the calibration exercise holds the exporter and new entrant lifecycles constant in the pre-reform steady states, differences in predictions across calibrations in Figure 2 can be attributed to features of the monopsony environment.

Monopsony gives a deeper response of the new exporter intensity profile to the liberalisation, as shown in Figure 2a.<sup>29</sup> Moreover, the full extent of this effect is only realised by relatively mature exporters in  $C_I$ . This contrasts with  $C_P$ , where the intensity changes are relatively uniform with respect to exporter age. The first year intensity response is 3.7pp in  $C_I$ , in contrast with 2.5pp in  $C_P$ , while the older age responses are 5.7pp and 5.3pp, respectively. The tariff reduction has no direct effect on a firm’s  $H$  sales in  $C_P$ . But with increasing marginal costs, exporting firms optimally lower their supply to  $H$  in  $C_I$ , thereby giving an amplified response in relative export sales (as discussed in Figure 1b). The age-biased effect in  $C_I$  follows from the predictions regarding exporter survival, depicted in Figure 2b — the probability of continuing to export from one period to the next rises by twice as much for first year exporters in  $C_I$  than it does in  $C_P$ . As a consequence, young exporters with the higher iceberg cost  $\tau_H^X$  tend to survive longer in the presence of monopsony.

Why does the survival rate for a first-year exporter rise from the tariff reduction so much

<sup>28</sup>Note that the steady-state lifecycle moments we discuss next are held constant across the calibrations in pre-reform steady state, by design. Consequently, we move straight away to discussion of the labour market environment’s impact on counterfactuals.

<sup>29</sup>For a given age of a firm’s export branch, this figure shows how much higher its average export intensity will be in the post-liberalisation lifecycle versus in the calibrated steady state. For instance,  $C_I$  predicts an export intensity for a first year exporter that is 3.7pp higher post-liberalisation (6.7% pre v.s. 10.4% post).

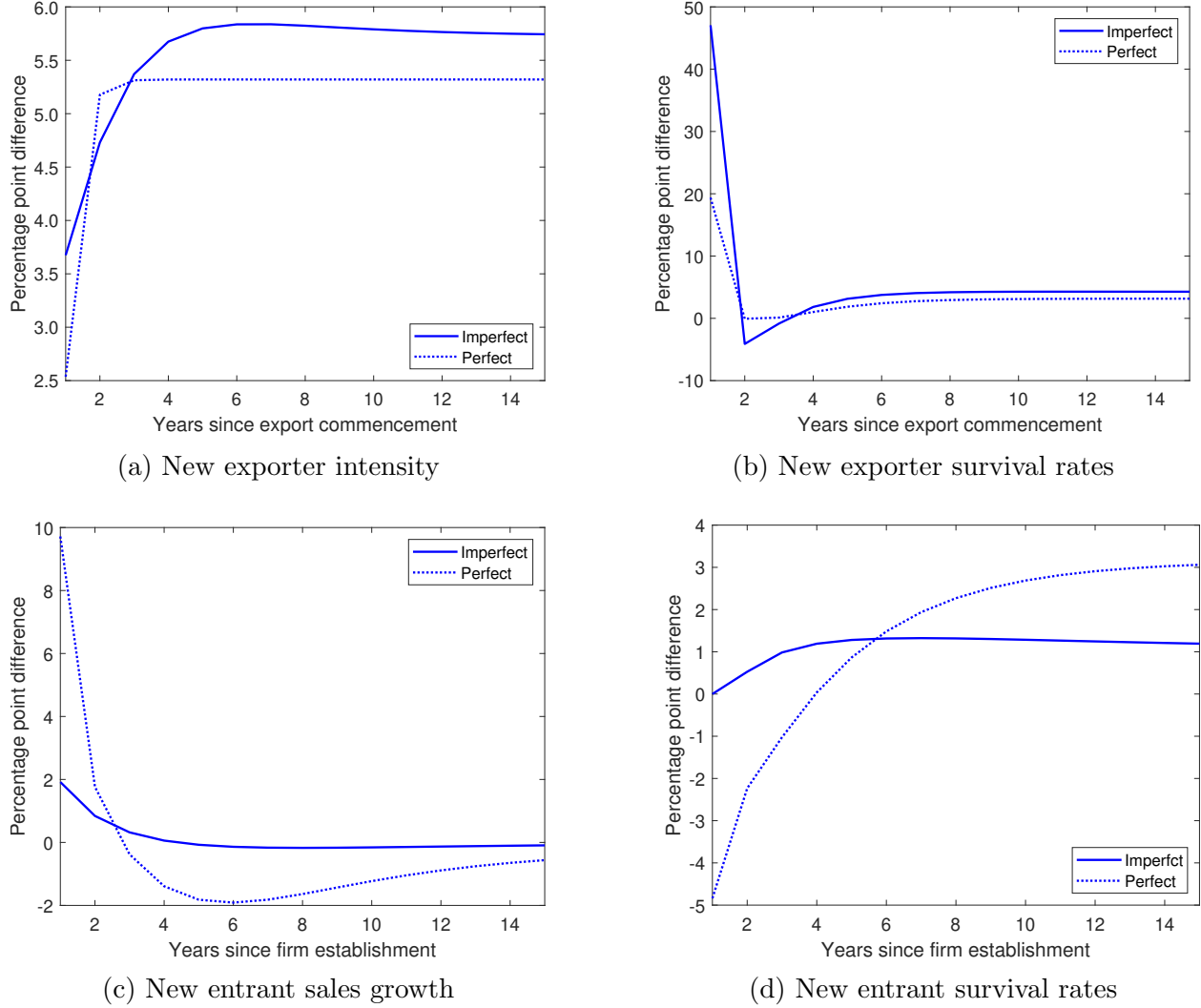


Figure 2: Firm expansion trajectories and survival rates. Cross-sectional comparisons across pre and post-liberalisation steady states. All variables are expressed as percentage point differences of the relevant rate in the post-liberalisation steady state from the pre-liberalisation steady state for the same given year since firm establishment/export commencement. All numbers are post-multiplication by 100.

more with monopsony power (as in Figure 2b)? Higher variable costs (wage bills due to USLS) associated with  $C_I$  yield lower calibrated sunk and period-by-period fixed costs than in  $C_P$ . Moreover the ratio of sunk/fixed costs is much lower in the former (1.22) than the latter (7.02). That is, it is relatively less expensive to establish and maintain an export branch in  $C_I$  than  $C_P$ . As such, marginal firms that upgrade from  $D$  pre-reform to  $X$  post-reform experience an increase in size of 9% in  $C_I$ , compared with 31% in  $C_P$ . Given that the trade elasticity is held constant across calibrations, this weaker firm-level response with monopsony therefore comes with a stronger extensive-margin reaction through higher survival.

Monopsony tends to dampen the volatility of the response of an entrant's lifecycle to the

tariff liberalisation. Figure 2c depicts the effect of the liberalisation episode on the expected growth rates of an entrant in its first 15 years of life.  $C_I$  predicts that entrants grow faster, for their first 7 years, than they would have at the same age pre-reform. A one-year-old firm grows around 2pp faster, roughly halving each year thereafter. This contrasts with  $C_P$ , where a one-year-old entrant grows 10pp faster post-reform, before in fact growing slower than it would pre-reform from age 3 onwards. Figure 2d paints a similar picture — the liberalisation gives new entrants a higher chance at survival at all ages in  $C_I$ , while they become more likely to exit post-reform until around age 4 in  $C_P$ . Two forces are responsible for these different predictions regarding entrants’ expected survival. Both calibrations predict entrants to upgrade faster to exporter status and remain there post-reform, as discussed in relation to Figure 2b. However, general equilibrium effects also make firms that retain domestic status more likely to exit. When averaging across firm establishments over each effect, the latter dominates in the early years of life under  $C_P$ , while the former prevails under  $C_I$ . The constant marginal cost environment in  $C_P$  offers a riskier portrait of entry — it allows for rapid scaling-up in the early years of operation, but comes coupled with firms ceasing operations at a much earlier age in expectation.

We summarise the cross-sectional results as follows. Trade liberalisation enhances outward market opportunities for firms, leading to an expansionary effect on the lifecycle of new exporters. This then spills over to affect the lifecycle of startups through continuation values, given the option value of becoming an exporter in the future. These effects are quantitatively significant in both  $C_I$  and  $C_P$ . However, with more prominence of variable relative to fixed costs,  $C_I$  gives a more conservative lifecycle response than  $C_P$  — new establishments exhibit weaker growth responses but receive a larger boost to their chance of survival. We now turn to examine the implications of these cross-sectional patterns for aggregation and welfare.

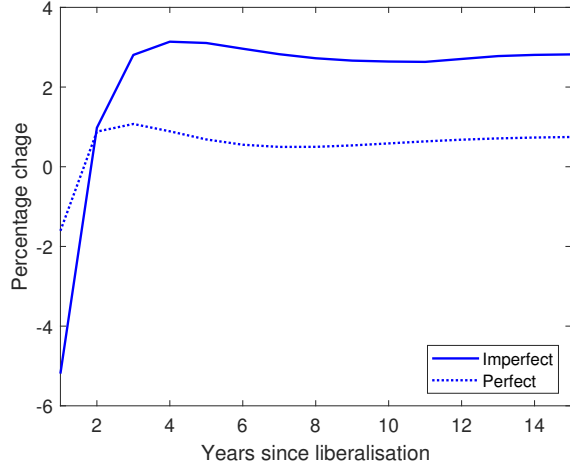
**Liberalisation episodes and aggregate variables.** Having explored how monopsony affects cross-sectional predictions, we are now in a position to discuss aggregation and transition dynamics. Figure 3 depicts the time paths of key macroeconomic aggregates, while Table 4 gives the percentage changes of variables across steady states.

Monopsony predicts a significantly stronger extensive-margin response of export activity. Our liberalisation episodes in  $C_I$  and  $C_P$  are designed to hold the long-run trade elasticity fixed across the two environments, as discussed in Section 6. Table 4 reports a decomposition of this elasticity into the extensive- and intensive-margin contributions.<sup>30</sup> Under imperfect labour markets, the extensive margin accounts for around 72% of the total trade response, compared with 67% under perfect labour markets. The difference reflects the fact that increasing marginal costs under monopsony dampen the ability of incumbent exporters to scale up production. Consequently, a larger share of the trade adjustment in  $C_I$  occurs through exporter creation, rather than incumbent expansion. Mechanically, this stronger extensive margin is facilitated by greater exporter survival, as depicted in Figure 2b.

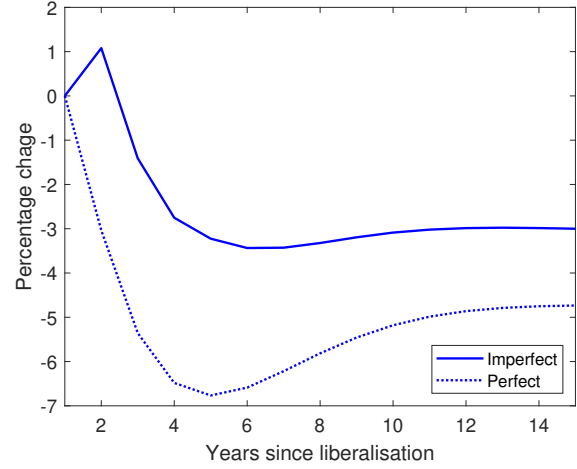
Empirical literature has documented the important role played by the extensive margin in

---

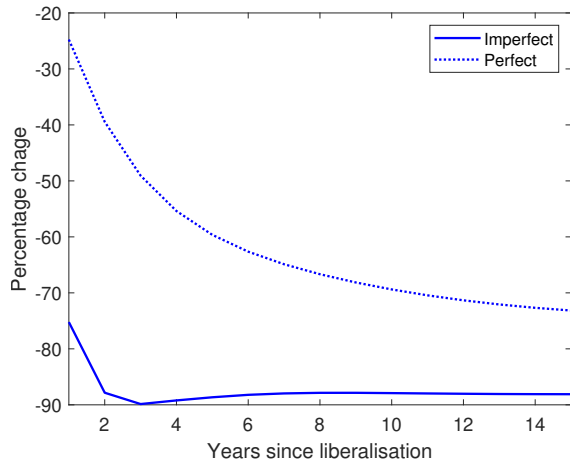
<sup>30</sup>We provide the exact expression for the intensive margin in Equation (A-142) in Appendix A.6.3. Specifically, we compute the intensive-margin response using the average physical iceberg cost  $\tau^X$  implied by the stationary exporter firm distribution and the pre-reform tariff rate. The extensive margin is then obtained residually so that the two components together give the overall trade elasticity.



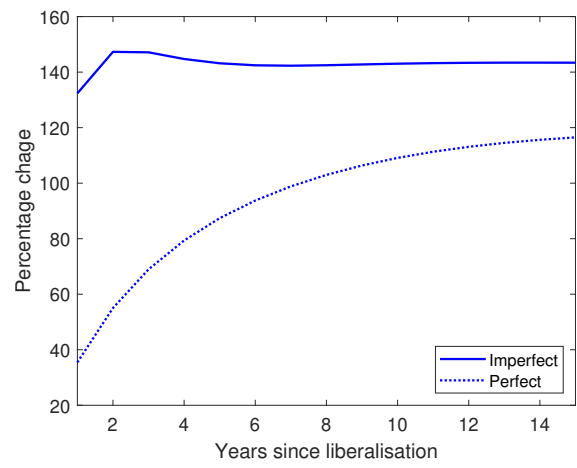
(a) Consumption



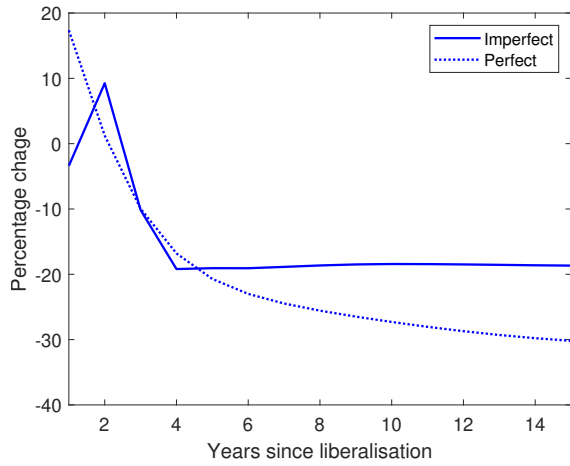
(b) Firms headquartered in  $H$  ( $\Omega_t$ )



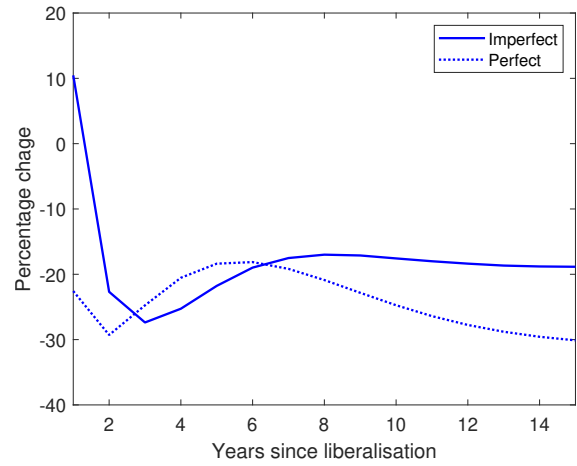
(c) Domestic ( $D$ ) firm mass ( $\Omega_t^D$ )



(d) Exporter ( $X$ ) firm mass ( $\Omega_t^X$ )



(e) Exiting ( $E$ ) firm mass ( $\Omega_t^E$ )



(f) Entering ( $T$ ) firm mass ( $\Omega_t^T$ )

Figure 3: Transition dynamics after tariff liberalisation. All variables are presented as percentage changes/deviations from initial calibrated steady state. All numbers are post-multiplication by 100.

	Trade		FDI	
	Imperfect	Perfect	Imperfect	Perfect
<i>Steady-state aggregates</i>				
Welfare	2.623	0.629	-0.721	0.000
Consumption	2.801	0.762	-0.700	0.010
Disutility	3.746	1.452	-0.589	0.078
Measure $D$ ( $\Omega^D$ )	-88.11	-74.79	-0.557	-1.471
Measure $X$ ( $\Omega^X$ )	143.0	119.6	-13.44	-11.09
Measure $M$ ( $\Omega^M$ )			61.27	52.46
Measure $T$ ( $\Omega^T$ )	-18.82	-31.48	-0.966	-1.332
Measure HQ ( $\Omega$ )	-3.150	-4.636	-1.328	-1.332
Measure $U$ ( $\Omega^U$ )	36.14	28.32	-1.721	-1.262
Profits	3.757	1.452	-0.509	0.344
Taxes	-100.0	-100.0	-20.43	-19.30
Wage index	3.122	1.208	-0.500	0.065
<i>Trade elasticity decomposition</i>				
Elasticity	12.98	12.30		
Intensive margin	28.17	32.52		
Extensive margin	71.83	67.48		
<i>Output decomposition</i>				
Final output	2.997	0.738	-0.672	0.000
Variety	7.712	6.233	-0.433	-0.317
Resources	-4.715	-5.495	-0.240	0.317
<i>Transition inclusive</i>				
Welfare	2.029	0.444	-0.531	0.007

Table 4: Changes in aggregate variables from bilateral liberalisation episodes. Unless stated otherwise, entries are percentage deviations from the initial steady state ( $\times 100$ ). Top panel *Steady-state aggregates* reports changes in steady-state aggregates. Second panel *Trade elasticity decomposition* presents steady-state trade elasticities and the contributions of the intensive and extensive margins; each component is normalised relative to the overall trade elasticity. Third panel *Output decomposition* decomposes the steady-state change in final aggregate output — variable  $A$  in Equation (6) — using  $A = (M^U)^{\frac{1}{\sigma-1}} B$ , where  $M^U$  is the mass of varieties in final goods aggregation (so  $(M^U)^{\frac{1}{\sigma-1}}$  corresponds to the “Variety” object in the table) and  $B$  denotes “Resources” — final output absent love-of-variety external scale economies. Entries in this panel are log deviations from the initial steady state ( $\times 100$ ), so that Variety and Resources sum to Final output. Bottom panel *Transition inclusive* reports lifetime consumption-equivalent welfare gains, appropriately discounted and including the transition.

trade flows. Using U.S. export data and distance as the trade-barrier measure, [Bernard, Redding and Schott \(2011\)](#) find that the negative relationship between U.S. exports and destination distance is fully accounted for by the extensive margin; [Lawless \(2010\)](#) presents similar evidence for U.S. exports and distance, as well as for other trade-cost proxies such as

language and infrastructure. [Foster \(2012\)](#) shows that trade liberalisation through Preferential Trade Agreements (PTAs) operates largely through the extensive margin. Using a large panel of countries for the 1962–2000 period, he finds that while PTAs increase trade along both margins, the majority of the resulting trade expansion reflects an increase in the variety of traded goods rather than higher volumes of existing products. [Dutt, Mihov and Zandt \(2013\)](#) find that common WTO membership increases trade flows at the extensive margin by about 25% while reducing them at the intensive margin by about 7%. [Baier, Bergstrand and Feng \(2014\)](#) provide systematic gravity-based evidence that trade liberalisation through economic integration agreements affects both intensive and extensive margins of trade, with economically meaningful responses along the extensive margin. Their results highlight the extensive margin as a central channel through which sustained reductions in trade costs translate into higher bilateral trade flows, particularly over longer horizons.

The firm cross-section converges much faster to the new steady state in  $C_I$  than  $C_P$ . The  $D$  and  $X$  masses have roughly converged within 4 years of the reform in  $C_I$ , contrasting with a gradual adjustment over the 15-year period in  $C_P$ . The more gradual intensive-margin adjustment with monopsony depicted in [Figure 2a](#) is complemented by more explosive dynamics in firm masses, as the economy adjusts to the time path of final-good demand. Quantitatively, the impact jump in the exporter mass is over three times larger in  $C_I$  than in  $C_P$ , constituting around 90% of the overall steady-state response in the former, while only 30% in the latter.

More action in the extensive margin response of the  $D$ - $X$  margin under monopsony also translates into more explosive entry and exit dynamics. Tariff liberalisation enhances new entrants’ potential for future profits via exporting, but also increases competitive pressure from imported goods — a trade-off that evolves over time. Following tariff reductions, dynamic models of this kind typically predict an immediate decline in entry and along the transition (e.g. [Alessandria, Choi and Ruhl, 2021b](#)); we replicate this result under  $C_P$ . In contrast, we overturn this pattern under  $C_I$  — the entry mass rises by 10% on impact, given firms’ surge in future prospects. Recall that, in our model setup, exporters commence operation with a one-year delay after upgrading. Once the new exporters from each country come online in the second period after liberalisation, upward pressure on factor prices and lower continuation values cause entry to drop in  $C_I$  relative to the steady state. The measure contracts for another period, before closing the gap with the old steady-state value by roughly one half. The shape of the entry response looks roughly similar qualitatively in  $C_P$  for around 7 years, before experiencing a second contraction downwards towards its new steady state. Exit is more volatile along the first few years of the transition in the presence of monopsony. High continuation values on impact cause a drop in exit in  $C_I$ , before rising for a year on account of additional entrants who are ejected from the market. Instead, exit rises for the first two years after reform in  $C_P$ , before eventually settling with a long run contraction that is nearly twice that in  $C_I$ .

Significantly different firm composition dynamics across  $C_I$  and  $C_P$  aggregate to drive starkly contrasting time paths for consumption, both in the short and long run. In [Figure 4](#), we decompose the different elements of the household’s budget constraint — labour income ( $WN$ ), profits gross of fixed costs ( $\Pi$ ), fixed cost expenditures ( $F$ ), and tax collections re-

distributed from the government ( $T$ ). The impact-year responses of most income sources are relatively similar across  $C_I$  and  $C_P$ ; the significantly larger drop in consumption under  $C_I$  stems from the fixed costs associated with new establishment creation. More investment in startups and exporters drives up wages, leading to labour income growth that lasts twice as long with monopsony, settling with a long-run effect nearly three times larger than  $C_P$ . Note that labour supply is distorted in both calibrations — in  $C_P$  due to product market markups, while  $C_I$  has the additional distortion of labour market markdowns. Consequently, since liberalisation increases labour supply, it alleviates this larger distortion with monopsony, yielding stronger welfare gains through labour income. The steady-state welfare change is around four times larger in  $C_I$  than  $C_P$ , while taking account of the transition path opens the gap up to nearly a factor of five. The faster cross-sectional adjustment on the firm-side of the economy also leads to faster consumption convergence with  $C_I$  — the ratio of lifetime to steady-state welfare gains is 77%, compared with 71% in  $C_P$ .

In Table 4, under *Output decomposition*, we report steady-state changes in final output, decomposed into contributions from love-of-variety scale economies and from a version of the final-good aggregator that abstracts from this effect. General equilibrium forces generate a decline in average output in both parameterisations, reflected in the leftward shift of the firm sales distribution shown in Figure 1d. However, under  $C_P$ , the more dispersed firm size distribution amplifies general equilibrium reallocations, resulting in disproportionately larger reductions in output among the most productive firms. As Table 4 shows, this leads to larger losses in aggregate resources (net of the love-of-variety effect) under perfect labour markets.<sup>31</sup> When combined with a weaker variety effect, this produces a significantly more muted overall final output response compared with  $C_I$ .

**Taking stock.** Monopsony delivers welfare gains from tariff liberalisation that are over four times larger than under perfect labour markets. It also amplifies both the short-run contraction and the long-run expansion during the transition. These differences stem from a stronger response of the extensive margin of exporting, on account of a weaker intensive margin, driven by monopsony’s impact on firm cost structures. These results paint an optimistic picture of monopsony’s effect on the efficacy of openness reforms. Motivated by our previous finding in the static model of Section 4 that FDI can lower welfare, we now turn to quantify the role of monopsony with respect to a key stage in a mature firm’s lifecycle: becoming a multinational.

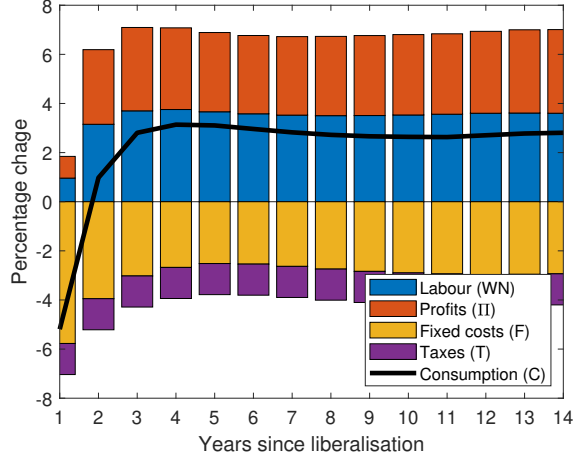
## 7.2 The Escape Effect and FDI Liberalisation

Similarly to Section 7.1, this section on FDI liberalisation builds up from studying firm decision rules in the steady state, before turning to quantify the effects of liberalisation on the cross-section and macroeconomic aggregates.

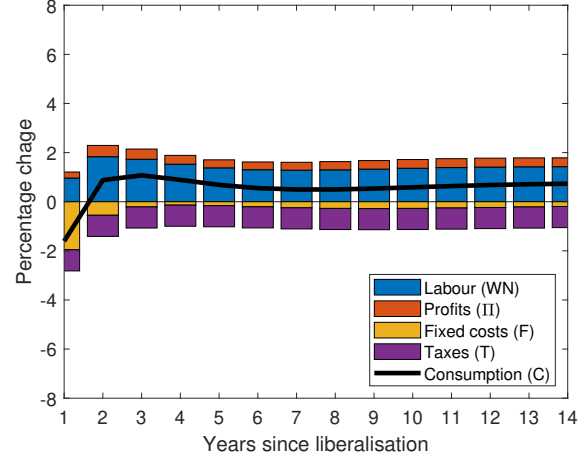
**Pre-reform steady state.** Previous discussion in Section 4 established that FDI allows firms to pay lower wages than operating as exporters. Here we relate this idea quantitatively

---

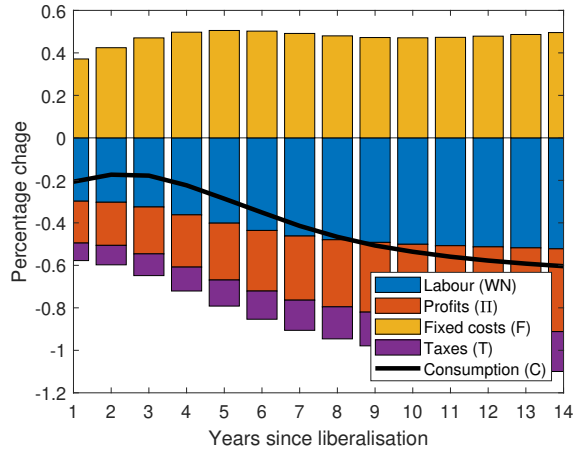
<sup>31</sup>The negative effect on resources follows since more exporting firms results in greater labour wastage through variable iceberg costs.



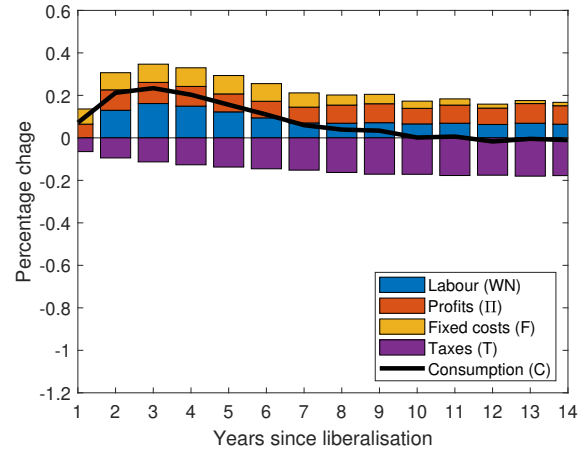
(a) Tariff reform with imperfect markets



(b) Tariff reform with perfect markets



(c) FDI reform with imperfect markets



(d) FDI reform with perfect markets

Figure 4: Consumption decomposition using resource constraint components. Note that the scale on the vertical axes are the same in 4a and 4b, as well as across 4c and 4d. Bars are expressed as absolute difference of the relevant component from its pre-reform steady-state value, expressed relative to steady-state pre-reform consumption. I.e. the sum of the bars for a given period give the percentage deviation of consumption (black line). All numbers are post-multiplication by 100.

to the firm lifecycle. Upgrading from  $X$  to  $M$  status (as opposed to moving from  $D$  to  $M$ ), as firms almost always do in the model and data, allows them to *escape* their locally-increasing marginal cost structure. Figures 1a and 1b convey all the information regarding how this part of the firm lifecycle affects wages and  $H$  output, respectively.<sup>32</sup> A firm upgrading from  $X$  to  $M$  with iceberg cost  $\tau^X$  will move back down from its elevated wage policy function to that of a domestic  $(z, D, \infty)$  firm, as it splits its production across two cost structures.

<sup>32</sup>Note that the slope and positioning of the policy functions in the FDI calibrations are very small quantitatively.

It continues to offer this menu of wages for its local workers, even when becoming a mature multinational (drawing  $\tau_L^M$ ).<sup>33</sup> A similar inference follows regarding its supply of goods to the  $H$  market — post-upgrade, the multinational firm’s supply policy function will coincide with that of a domestic firm. That is, while  $q(z, X, \tau_L^X) < q(z, X, \tau_H^X) < q(z, D, \infty)$ , firms with multinational status are such that  $q(z, M, \tau^M) = q(z, D, \infty)$  for  $\tau^M \in \{\tau_L^M, \tau_H^M\}$  in  $C_I^{FDI}$ . Once a firm undertakes horizontal FDI, its production for  $H$  and  $F$  is completely segmented, and its employment choices in each country are independent.

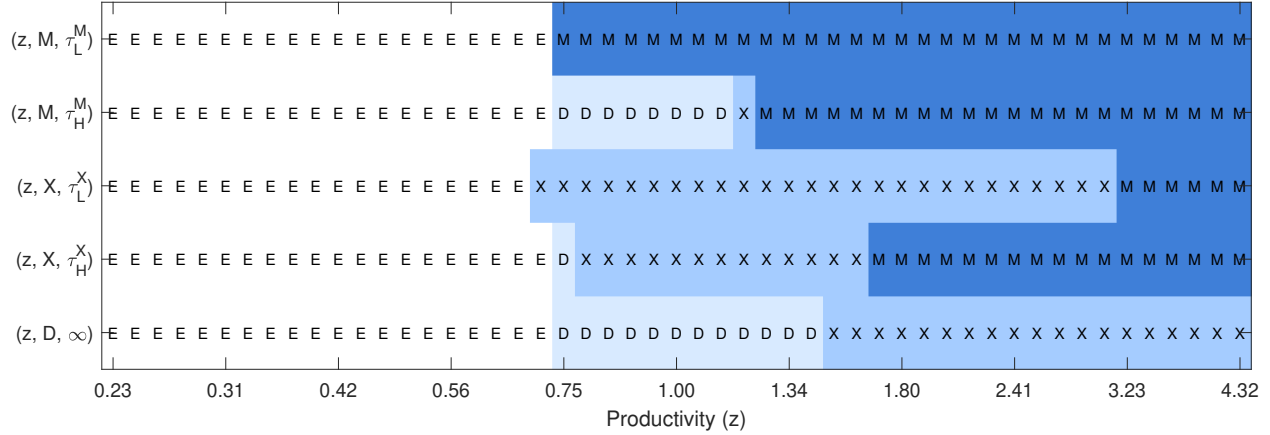
Figure 5 illustrates the discrete-choice decision rules of firms of all phases of the lifecycle for  $C_I^{FDI}$  (Figure 5a) and  $C_P^{FDI}$  (Figure 5b).<sup>34</sup> There are two main points to note regarding multinational activity and monopsony. In both calibrations, the bulk of the state space transitioning from  $X$  to  $M$  status is coming from exporters with the lower export-specific productivity (from the row marked  $(z, X, \tau_L^X)$ ). A higher export-specific productivity (i.e.,  $\tau_L^X < \tau_H^X$ ) implies that a higher fundamental productivity  $z$  is required to make it worthwhile to pay the FDI sunk cost, given the low initial FDI efficiency,  $\tau_H^M$ . The option to escape in  $C_I^{FDI}$  is so valuable that some exporters with the higher export efficiency (row  $(z, X, \tau_L^X)$ ) will choose to upgrade. This is the first main contrast involving monopsony and FDI: in  $C_P^{FDI}$ , only exporters with  $\tau_H^X$  upgrade to become multinationals — none with  $\tau_L^X$  do so. The second main point to note concerns the opposite event — downgrading from multinational status. Given that fixed costs play a larger role in firm profit structures with perfect labour markets, some multinationals with the higher specific efficiency (row  $(z, M, \tau_L^M)$ ) will downgrade from  $M$  to  $D$ . These same firms will retain FDI status in  $C_I^{FDI}$ .

**Liberalisation episodes and the firm cross-section.** We now implement a bilateral reduction in the outward FDI tax from its calibrated value to zero in both  $C_I^{FDI}$  and  $C_P^{FDI}$ . Figure 6 shows changes in firm lifecycle trajectories, similarly to Figure 2, comparing across post-reform and pre-reform steady states in each parameterisation.

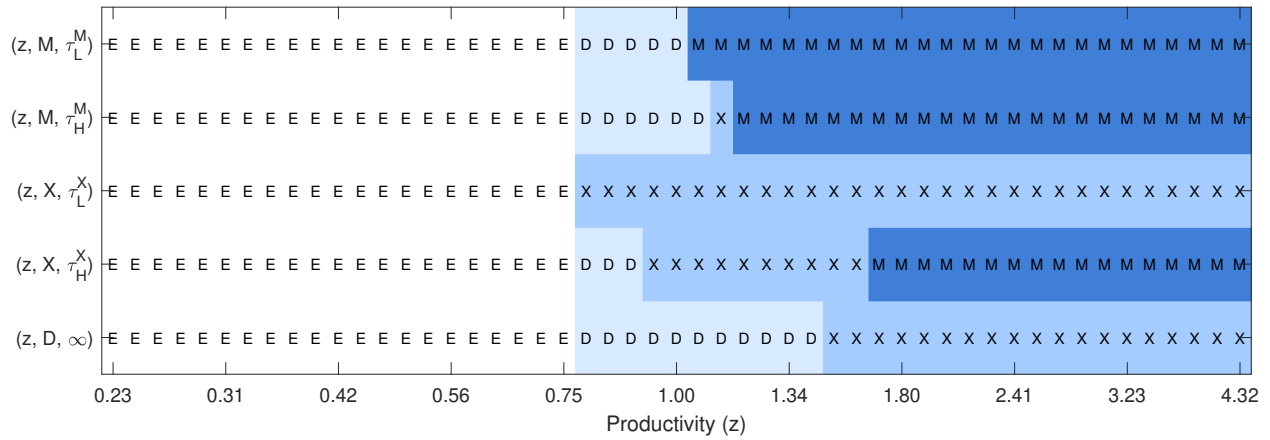
Figures 6a and 6b show the impact of the reform across the first 15 years of a new multinational’s life in terms of FDI intensity and survival probability, respectively. In interpreting these results, recall the discussion from Section 3.3.1: in contrast to the tariff, the FDI tax has no partial-equilibrium impact on the intensive margin. All adjustments from the FDI tax ensue as a result of selection and general equilibrium effects only. The reform generally leads more firms to retain the status in both  $C_I^{FDI}$  and  $C_P^{FDI}$ . This translates into a drop in the average intensity profile at most points in the lifecycle, since less efficient multinationals (with  $\tau_H^M$ ) survive longer. Guided by discussion of Figure 5, the gains in firm value when upgrading from  $X$  to  $M$  in  $C_I^{FDI}$  are front-loaded — the option to escape gives firms immediate variable cost savings as soon as their subsidiary comes online. In contrast, the gains

<sup>33</sup>This does not imply that multinational firms pay lower wages on average in  $H$  in our calibration, as these are merely the wage policy functions. Due to selection effects, multinational firms are, on average, more productive than exporters. As a result, they produce more for  $H$  and pay correspondingly higher wages. Quantitatively, in  $C_I^{FDI}$ , the average  $H$  wage paid by a multinational firm is approximately 20% higher than that of an exporter.

<sup>34</sup>Bear in mind that our calibration procedure holds the *average* entrant, exporter, and multinational lifecycles constant; differences in the firm distribution and policy functions across the two calibrations average out. In fact, the firm sales distribution looks similar qualitatively and quantitatively with FDI as in Figure 1d, as such, we omit it from the writeup.



(a) Imperfect labour markets

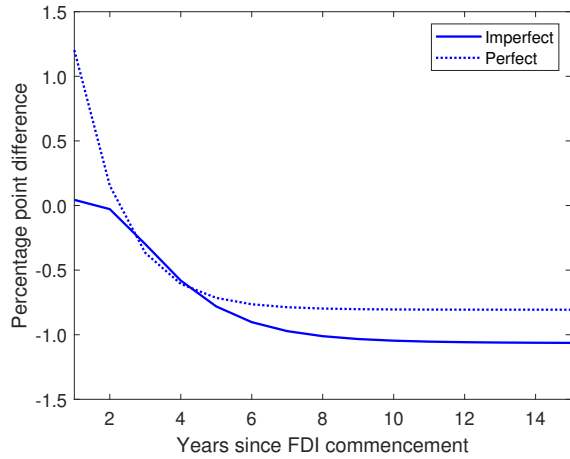


(b) Perfect labour markets

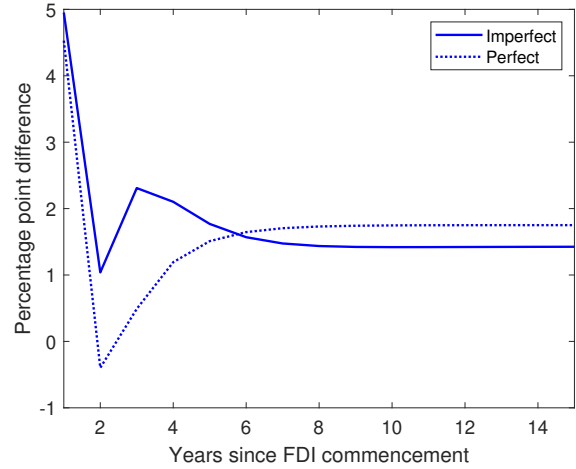
Figure 5: Firm status policy functions. Choice  $E$  represents exit,  $D$  for domestic,  $X$  for exporter and  $M$  for multinational. Horizontal axis displays firm productivity while vertical grids give current firm state vector  $(z, s, \tau)$ . Note that all policy functions are drawn with a fixed value of the stochastic sunk cost draws.

are more back-loaded in  $C_P^{FDI}$ : firms realise the biggest savings when drawing  $\tau_L^M$  later in their affiliate's life. This leads to a noteworthy contrast in Figure 6b — while the affiliate survival probability rises post-reform at all ages in  $C_I^{FDI}$ , it actually drops temporarily for second year affiliates in  $C_P^{FDI}$ . Firms upgrade to  $M$  status in  $C_P^{FDI}$  with the hope of quickly drawing the lower iceberg shock  $\tau_L^M$ ; some of those who do not realise this shock immediately will close their subsidiary. Following this cleansing effect, older affiliates stand a stronger chance of survival post-reform in  $C_P^{FDI}$  than in  $C_I^{FDI}$ , also driving a smaller decline in the FDI sales intensity for mature subsidiaries in the former.

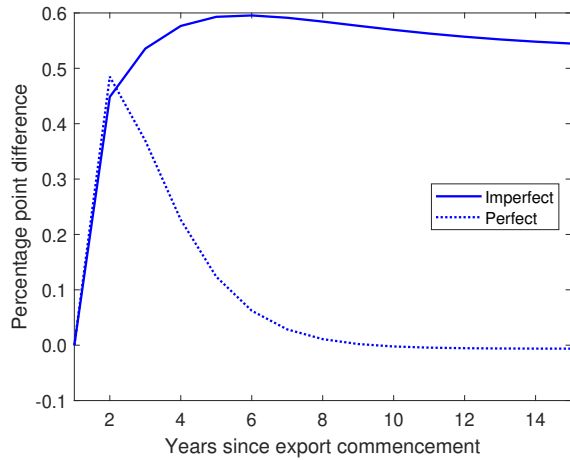
Differences in the multinational lifecycle counterfactual response under monopsony spill over to more muted lifecycle reactions of the statuses underneath — new exporters and entrants. Figure 6d shows that, while both  $C_I^{FDI}$  and  $C_P^{FDI}$  generally predict more churn of exporters,



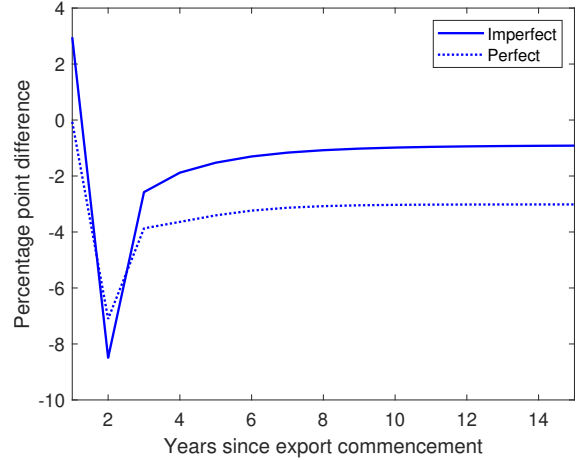
(a) New multinational intensity



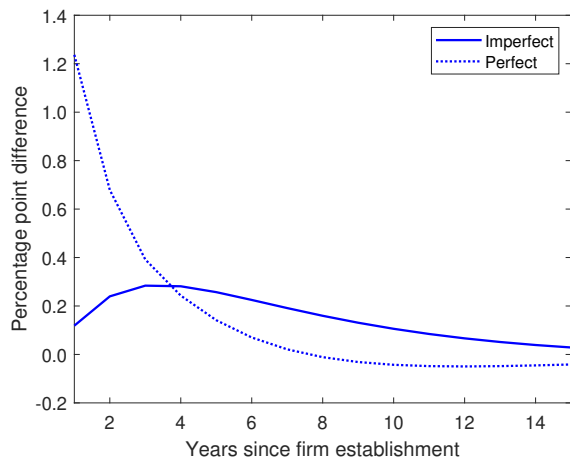
(b) New multinational survival rates



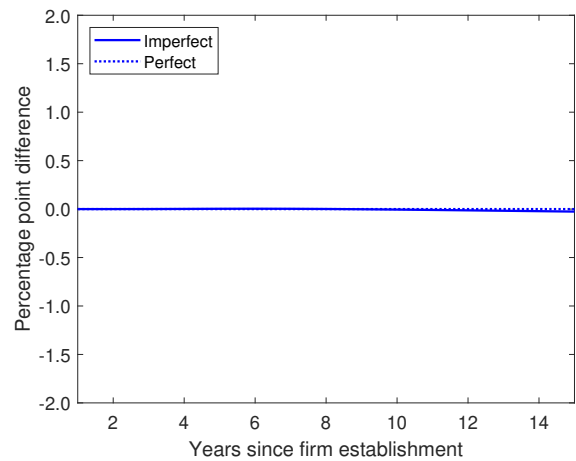
(c) New exporter intensity



(d) New exporter survival rates



(e) New entrant sales growth



(f) New entrant survival rates

Figure 6: Firm expansion trajectories and survival rates. Cross-sectional comparisons across pre and post-liberalisation steady states. All variables are expressed as percentage point differences of the relevant rate in the post-liberalisation steady state from the pre-liberalisation steady state for the same given year since firm establishment/export commencement. All numbers are post-multiplication by 100.

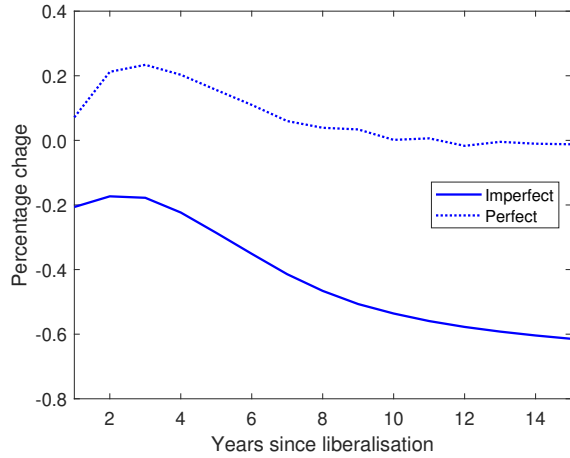
the effect is quantitatively stronger in the latter. For most exporter ages, the probability of leaving the status, either by upgrading to  $M$  or downgrading to  $D$ , rises by significantly more in  $C_P^{FDI}$ . This translates into markedly different new exporter intensity trajectories, depicted in Figure 6c. The liberalisation gives a 0.5pp increase in the export intensity of new exporters aged two in both  $C_I^{FDI}$  and  $C_P^{FDI}$ . The change remains flat thereafter in  $C_I^{FDI}$ , whereas it gradually declines in  $C_P^{FDI}$ . While startup survival probabilities are relatively unaffected in both parameterisations,  $C_P^{FDI}$  predicts a significantly faster sales growth response in the early years of operation, before giving slower growth from age six onwards. These results for statuses below  $M$  have a similar flavour to those for new entrants in the context of the tariff reform in Section 7.1. A higher option value for being a multinational in the future impacts the continuation values of new exporters and entrants. This interacts with the features of the environment we study, which generally yields a riskier portrait of new establishment/export branch creation in  $C_P^{FDI}$ .

**Liberalisation episodes and aggregate variables.** How do these cross-sectional changes impact the transition of aggregates? We present the transition paths in Figure 7 and steady-state results in Table 4.

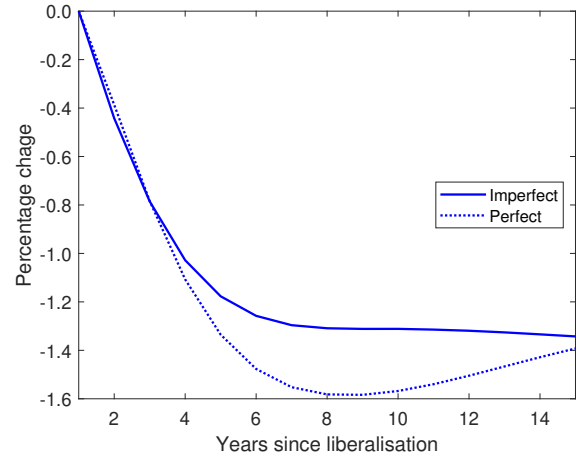
The long-run extensive margin of multinational activity is stronger in the presence of monopsony. Figure 7e shows that the mass  $\Omega_t^M$  follows a similar trajectory across  $C_I^{FDI}$  and  $C_P^{FDI}$  for the first five years following the reform, before they depart and the former settles at 61% higher than the initial steady state, compared with 53% in the latter. The long-run change in the mass of firms upgrading from  $X$  to  $M$  is similar across the two parameterisations (not shown in figure). Instead, the difference in the steady-state  $M$  mass change is driven by the survival margin discussed in Figure 6b. Although older affiliates get a stronger survival boost with the reform in  $C_P^{FDI}$  than in  $C_I^{FDI}$ , the lower survival trajectory in the early years of life leads to more firm turnover, giving a higher measure of exiting multinationals each period.

Firms of lower statuses face tougher competition from multinationals after FDI liberalisation in both  $C_I^{FDI}$  and  $C_P^{FDI}$ , resulting in a decline in the mass of active firms in both cases. Guided by the discussion in Section 4,  $C_P^{FDI}$  gives a parameterisation where FDI is a positive for welfare, which is not the case in the presence of monopsony. In  $C_I^{FDI}$ , as more firms attain multinational status and shift production abroad, domestic wages decline due to reduced labour demand. This lowers aggregate household income and thus depresses demand for the remaining domestic firms, particularly intermediate-productivity firms that cannot profitably undertake FDI. The result is an inefficient contraction of the domestic firm mass and a drop in total domestic supply. In contrast, under perfect labour markets,  $C_P^{FDI}$ , multinationals expand without depressing wages or aggregate income, so the exit of intermediate-productivity firms reflects efficient reallocation rather than a distortion-induced contraction.

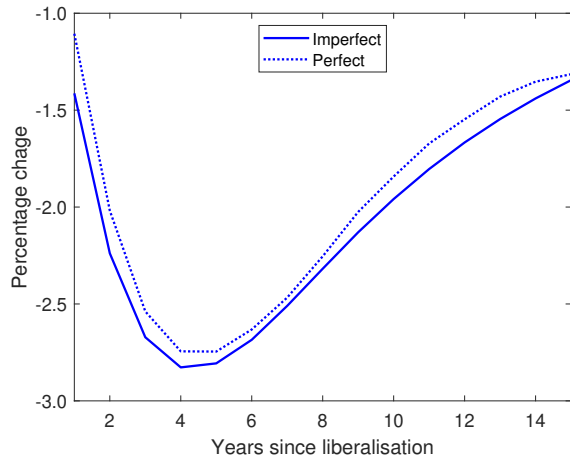
Consumption's transition path to the reform in  $C_I^{FDI}$  can be roughly seen as that in  $C_P^{FDI}$  translated downwards. The difference is quantitatively meaningful — while the 1% tax reduction gives roughly a zero long-run consumption response in  $C_P^{FDI}$ , it gives a substantial 0.7% loss in  $C_I^{FDI}$ . All of the consumption benefits of the reform in  $C_P^{FDI}$  are realised only along the transition. Figures 4c and 4d illustrate substantially different contributions from



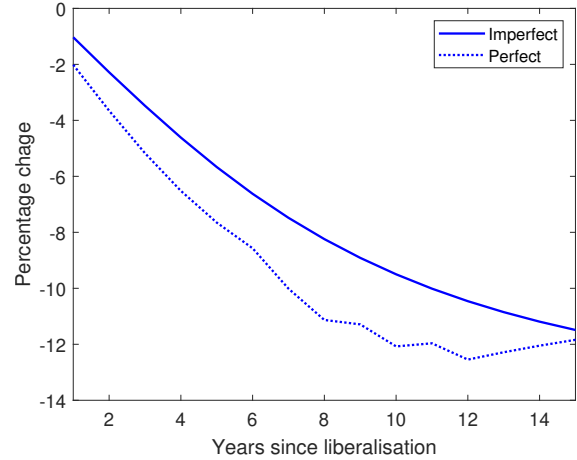
(a) Consumption



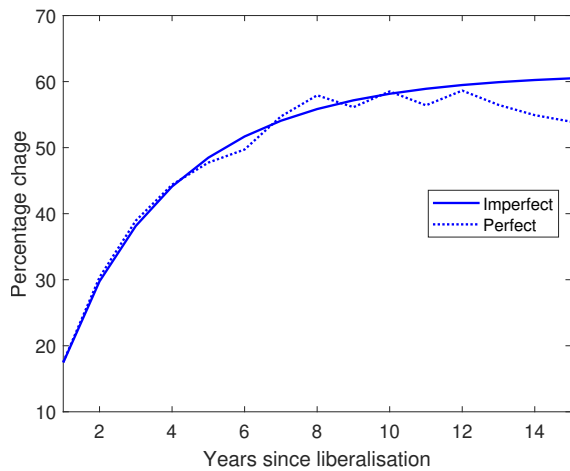
(b) Firms headquartered in  $H$  ( $\Omega_t$ )



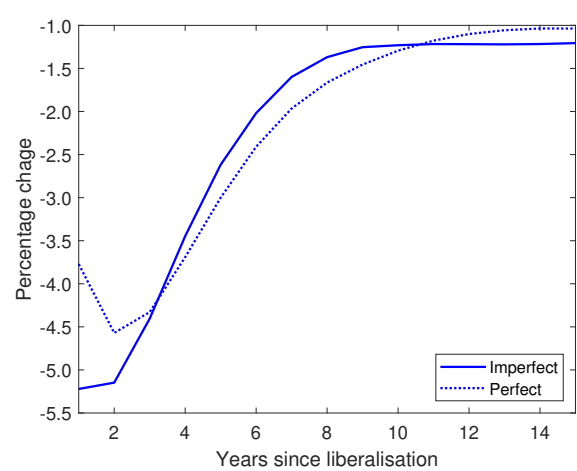
(c) Domestic ( $D$ ) firm mass ( $\Omega_t^D$ )



(d) Exporter ( $X$ ) firm mass ( $\Omega_t^X$ )



(e) Multinational ( $M$ ) firm mass ( $\Omega_t^M$ )



(f) Entering ( $T$ ) firm mass ( $\Omega_t^T$ )

Figure 7: Transition dynamics after FDI liberalisation. All variables are presented as percentage changes/deviations from initial calibrated steady state. All numbers are post-multiplication by 100.

sources of income in driving the consumption transition. The drop in new entry and export branches releases fixed costs in both calibrations, being a roughly constant offsetting force in  $C_I^{FDI}$  and declining in prominence in  $C_P^{FDI}$ . This effect, combined with declining labour and tax contributions over time, drives the zero steady-state effect in  $C_P^{FDI}$ . Higher profits also ensue with perfect labour markets, with a relatively constant impact over time. In contrast, more FDI leads to lower average wages in  $C_I^{FDI}$ , an effect that grows over time. As  $M$  firms pay lower wages, this makes it more difficult for domestic and exporting firms to attract workers, leading to a negative contribution from profits that strengthens with time. Note also that Table 4 quantifies the movements in final aggregate output attributable to love of variety. For both parameterisations, the variety contribution moves in the opposite direction to tariff reform. Fewer varieties ensue, as more resources move towards the sunk/fixed costs of multinational activity. Falling wages drive lower resource utilisation with imperfect markets, following from anti-selection effects, as can be seen with the negative profit response. This contrasts with perfect markets where a positive resource effect offsets the loss from varieties.

These aggregate results paint a different picture of monopsony’s impact on the efficacy of FDI liberalisation episodes compared with tariff liberalisation. The common ground is that the monopsony friction is a quantitatively powerful source of amplification in the aggregate for both types of reforms, coming through the environment’s impact on firms’ choices and lifecycle responses. The short- and long-run effects can be very different quantitatively, necessitating the use of our dynamic analysis. Our results suggest a positive normative association for liberalising exports in markets with monopsony on the input side, but the opposite with regard to horizontal FDI. Reforms to trade and tax policy are usually enacted simultaneously, implying that monopsony’s net effect on the associated openness counterfactual predictions will generally be ambiguous.<sup>35</sup> This work has aimed to highlight the quantitative significance of the monopsony mechanism — we leave further quantification of particular policy episodes with a multi-sector and multi-country structure as an avenue for future research.

## 8 Robustness and Extensions

### 8.1 Parameter Robustness

Table A-1 presents comparative statics on our steady-state results for some key parameters. In particular, we consider a higher value of  $\theta$ ,  $\theta = 6.5$ , taken from estimates in Lamadon, Mogstad and Setzler (2022); set  $\phi = 0$ , consistent with the discussion in the simplified framework of Section 4; and consider lower values of the elasticity of substitution across varieties,  $\sigma$ . We change the parameter value in question and then re-calibrate the model each time to match the same targets as in Table 3 for the parameterisations without FDI, then explore the effect of the bilateral tariff reduction. For expositional ease, we continue to refer each time to the monopsonistic labour market structure as  $C_I$  and the perfect labour market parameterisation as  $C_P$ .

As the elasticity of firm-level labour supply rises to  $\theta = 6.5$ , the parameterisation moves

---

<sup>35</sup>For example, the U.S.-China trade war and the Tax Cuts and Jobs Act in 2018.

closer to the limit case of  $\theta \rightarrow \infty$ , which constitutes  $C_P$ . Given that this lowers the slope of the firms' increasing marginal cost structures, it brings down the monopsonistic welfare gains, from 2.6% to 1.5%; notice that the latter gains are still around 2.5 times larger than the prediction with perfect labour markets of 0.6%. This value of  $\theta = 6.5$  is from the top of the set of estimates of this elasticity from the empirical literature, meaning these differences from the perfect markets counterfactual can be thought of as a conservative lower-bound. Recall that our re-calibration process holds the long-run trade elasticity constant across parameterisations. As such, robustness on the other parameters  $\phi$  and  $\sigma$  tend to either leave the differences across  $C_I$  and  $C_P$  relatively unaffected, or to slightly increase monopsony's effect.

## 8.2 Decomposing WMD and USLS

Our monopsony environment features an upward-sloping labour supply (USLS) and wage markdowns (WMD). Table A-2 decomposes the long-run effects of these two features in driving aggregate quantitative long-run differences of the two environments. In the context without FDI, calibration  $C_U$  retains USLS while imposing that firms do not internalise their wage-setting power — that is, they behave myopically. We do this also for the calibrations with FDI; as usual, all scenarios are re-calibrated before performing counterfactual exercises. Comparing  $C_I$  to  $C_U$  isolates the effect of WMD, while comparing  $C_U$  to  $C_P$  captures the impact of USLS.

WMD have no bearing on the slope of firms' marginal cost structures, but rather affect their level of profits. USLS drives most of the difference in quantitative predictions in the case of trade liberalisation. A slightly weaker wage effect follows, given a stronger drop in the measure of entrants in  $C_I$  than  $C_U$ . These small differences likely reflect the relatively low export intensity matched in the calibration (15%). Exporting firms earn incremental profits over domestic firms, which are influenced by WMD, though these gains are relatively modest. In contrast, multinationals have a foreign sales intensity of around 30%, while also being significantly larger than exporters in terms of size. A firm that exports prior to the FDI liberalisation sits a long way along its marginal cost structure in  $C_U^{FDI}$ . As such, the elasticity of the option to escape is very powerful when the outward tax is removed in the counterfactual. This leads to an amplified extensive-margin response and correspondingly larger losses in wages and welfare — welfare declines by 3.2% in  $C_U^{FDI}$  and by 0.7% in  $C_I^{FDI}$ .

## 8.3 Risk Averse Households

Our model assumes risk neutrality on the part of workers. We briefly examine how the quantitative results change when generalising to risk-averse preferences. Appendix C.3 details the modifications to the model, which effectively introduce an income effect that counteracts the labour supply expansion induced by higher wages.<sup>36</sup> Figure A-1 depicts how the lifecycle responses of exporters and new entrants to tariff liberalisation are affected under risk aversion. With the weaker expansion in aggregate labour supply, fewer resources are available to accommodate the creation of new exporters and startups. This drives a deeper expansion in

---

<sup>36</sup>The labour-supply disutility aggregate in Table A-3 illustrates this mechanism.

the export intensity for both  $C_I^{RA}$  and  $C_P^{RA}$  than their counterparts with risk neutrality. This also comes with dampened entrant sales growth responses, as well as a stronger tendency for survival. Table A-3 shows that the labour supply expansion in  $C_I^{RA}$  is 1/5 of that in  $C_I$ , giving a smaller absolute difference in welfare gains of  $C_I^{RA}$  from  $C_P^{RA}$ . However the basic inference regarding relative welfare differences is robust — the gains with monopsony are approximately three times larger than with perfect labour markets.

## 8.4 Offshoring

A key issue in recent years’ policy discussions regarding labour markets relates to offshoring of the production process. To account for this in our model, we consider an asymmetric country parameterisation, where  $F$  has a larger labour supply elasticity  $\theta \rightarrow \infty$  while  $H$  production retains  $\theta = 1.88$ . Firms in  $H$  have the option to undertake all of their production in  $F$ , then ship some of their goods back to their home market. Appendix C.4 provides more details and results.

Figure A-2 plots the steady-state output policy functions of  $H$ -headquartered firms, showing their output supply to each market across different statuses. When a firm offshores its production, it incurs additional iceberg efficiency losses on output destined for  $H$  households but gains access to a cost structure with constant marginal cost. As a result, firms with intermediate levels of desired scale likely have limited incentive to offshore. Only firms that have considerable desired output scale have incentive to escape the increasing  $H$  marginal cost structure. This agrees with empirical findings that only the most productive firms have incentive to offshore (e.g. see Zlate, 2016; Kurz, 2006). Moreover, this exercise highlights our main channel through which monopsony affects the supply side of the economy — by influencing firms’ organisational form choices based on cost structure. Incorporating vertical FDI into a dynamic model of firms and monopsony, using an asymmetric parameterisation, presents a promising avenue for future research. Such an approach would also offer a framework to interpret compelling empirical patterns concerning the presence of multinationals and local labour market outcomes in developing countries.<sup>37</sup>

## 9 Conclusion

Motivated by abundant empirical evidence that firms operate in monopsonistic labour markets, we build the first dynamic laboratory for quantifying monopsony’s impact on counterfactual predictions. Our model extends the firm dynamics and open economy literature to generate rich predictions about how monopsony alters firm cost structures and behaviour.

We carefully calibrate the model to micro moments related to firm growth and foreign engagement, then simulate bilateral trade and FDI liberalisation episodes. Monopsony emerges as a quantitatively significant friction shaping aggregate adjustment — it dampens incumbent firm expansion while amplifying entry and new exporter creation along the transition path. These forces magnify the welfare gains from tariff reductions but also generate poten-

---

<sup>37</sup>See, for example, Alfaro-Ureña et al. (2021) for empirical evidence using microdata on Western multinationals in Costa Rica.

tial welfare losses from outward FDI liberalisation, as firms exploit multinational status to escape rising domestic labour costs.

These findings suggest that when trade and international investment policy are used to bolster domestic labour market activity and worker welfare, monopsony frictions deserve central consideration. Within the purview of exporting and FDI as substitutes, the stage of the firm lifecycle that a policy primarily affects is important. Monopsony amplifies the effectiveness of trade policy in stimulating young firm activity and strengthening domestic labour markets, but undermines the effectiveness of FDI liberalisation, which primarily benefits large, mature, and profitable firms later in the lifecycle. Important extensions of this study include allowing for variable wage markdowns to explore how liberalisation affects labour market competition, as well as introducing a richer multi-country setting with offshoring activity. These results provide a foundation for more dynamic analyses of labour market monopsony and global production decisions, while also informing policy debates surrounding multilateralism and the role of firms in the global economy.

## References

- Alessandria, George and Horag Choi**, “Do sunk costs of exporting matter for net export dynamics?,” *The Quarterly Journal of Economics*, 2007, 122 (1), 289–336.
- **and –**, “Do falling iceberg costs explain recent US export growth?,” *Journal of International Economics*, 2014, 94 (2), 311–325.
- **and –**, “Establishment heterogeneity, exporter dynamics, and the effects of trade liberalization,” *Journal of International Economics*, 2014, 94 (2), 207–223.
- **, Costas Arkolakis, and Kim J Ruhl**, “Firm dynamics and trade,” *Annual Review of Economics*, 2021, 13 (1), 253–280.
- **, Horag Choi, and Kim J Ruhl**, “Trade adjustment dynamics and the welfare gains from trade,” *Journal of International Economics*, 2021, 131, 103458.
- Alfaro-Ureña, Alonso, Isabela Manelici, and Jose P Vasquez**, “The effects of multinationals on workers: Evidence from Costa Rican microdata,” Technical Report, Working paper 2021.
- Anderson, Simon P, Andre De Palma, and Jacques-Francois Thisse**, *Discrete choice theory of product differentiation*, MIT press, 1992.
- Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple**, “Innovation and production in the global economy,” *American Economic Review*, 2018, 108 (8), 2128–73.
- Atkeson, Andrew and Ariel Tomas Burstein**, “Innovation, firm dynamics, and international trade,” *Journal of political economy*, 2010, 118 (3), 433–484.
- Baier, Scott L., Jeffrey H. Bergstrand, and Michael Feng**, “Economic integration

agreements and the margins of international trade,” *Journal of International Economics*, 2014, *93* (2), 339–350.

**Bendetti-Fasil, Cristiana, Giammario Impullitti, Omar Licandro, Petr Sedláček, and Adam Spencer**, *Heterogeneous firms, growth and the long shadows of business cycles*, Centre for Finance, Credit and Macroeconomics, School of Economics . . . , 2024.

**Berger, David, Kyle Herkenhoff, and Simon Mongey**, “Labor Market Power,” *American Economic Review*, 2022, (1147).

– , – , and – , “Minimum Wages, Efficiency, and Welfare,” *Econometrica*, 2025, *93* (1), 265–301.

**Bernard, Andrew B, J Bradford Jensen, and Peter K Schott**, “Importers, exporters and multinationals: a portrait of firms in the US that trade goods,” in “Producer dynamics: New evidence from micro data,” University of Chicago Press, 2009, pp. 513–552.

**Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott**, “Multiproduct Firms and Trade Liberalization \*,” *Quarterly Journal of Economics*, 07 2011, *126* (3), 1271–1318.

**Bilbiie, Florin O, Fabio Ghironi, and Marc J Melitz**, “Monopoly power and endogenous product variety: Distortions and remedies,” *American Economic Journal: Macroeconomics*, 2019, *11* (4), 140–174.

**Boehm, Christoph E, Aaron Flaaen, and Nitya Pandalai-Nayar**, “Multinationals, offshoring, and the decline of US manufacturing,” *Journal of International Economics*, 2020, *127*, 103391.

**Brooks, Wyatt and Alessandro Dovis**, “Credit market frictions and trade liberalizations,” *Journal of Monetary Economics*, 2020, *111*, 32–47.

**Caliendo, Lorenzo and Fernando Parro**, “Estimates of the Trade and Welfare Effects of NAFTA,” *The Review of Economic Studies*, 2015, *82* (1), 1–44.

**Cooper, Russell W and John C Haltiwanger**, “On the nature of capital adjustment costs,” *The Review of Economic Studies*, 2006, *73* (3), 611–633.

**Corbae, Dean and Pablo D’Erasmus**, “Reorganization or liquidation: Bankruptcy choice and firm dynamics,” *The Review of Economic Studies*, 2021, *88* (5), 2239–2274.

**Costinot, Arnaud and Andrés Rodríguez-Clare**, “Trade theory with numbers: Quantifying the consequences of globalization,” in “Handbook of international economics,” Vol. 4, Elsevier, 2014, pp. 197–261.

**Dhyne, Emmanuel, Ayumu Ken Kikkawa, Toshiaki Komatsu, Magne Mogstad, and Felix Tintelnot**, “Foreign Demand Shocks to Production Networks: Firm Responses and Worker Impacts,” Working Paper 30447, National Bureau of Economic Research September 2022.

- Dutt, Pushan, Ilian Mihov, and Timothy Van Zandt**, “The Effect of WTO on the Extensive and the Intensive Margins of Trade,” *Journal of International Economics*, 2013, *91* (2), 204–219.
- Egger, Hartmut, Udo Kreickemeier, Christoph Moser, and Jens Wrona**, “Exporting and Offshoring with Monopsonistic Competition,” *The Economic Journal*, May 2022, *132* (644), 1449–1488.
- Fillat, José L and Stefania Garetto**, “Risk, returns, and multinational production,” *The Quarterly Journal of Economics*, 2015, *130* (4), 2027–2073.
- Foster, Neil**, “Preferential Trade Agreements and the Margins of Imports,” *Open Economies Review*, 2012, *23* (4), 869–889.
- Garetto, Stefania, Lindsay Oldenski, and Natalia Ramondo**, “Multinational expansion in time and space,” Technical Report, National Bureau of Economic Research 2019.
- Gumpert, Anna, Haishi Li, Andreas Moxnes, Natalia Ramondo, and Felix Tintelnot**, “The life-cycle dynamics of exporters and multinational firms,” *Journal of International Economics*, 2020, *126*, 103343.
- Helpman, Elhanan, Marc J Melitz, and Stephen R Yeaple**, “Export versus FDI with heterogeneous firms,” *American economic review*, 2004, *94* (1), 300–316.
- Hopenhayn, Hugo and Richard Rogerson**, “Job turnover and policy evaluation: A general equilibrium analysis,” *Journal of political Economy*, 1993, *101* (5), 915–938.
- Jha, Priyaranjan and Antonio Rodriguez-Lopez**, “Monopsonistic labor markets and international trade,” *European Economic Review*, 2021, *140*, 103939.
- and –, “Minimum Wage and Employer Variety,” 2022.
- Kurz, Christopher Johann**, “Outstanding outsourcers: A firm-and plant-level analysis of production sharing,” 2006.
- Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler**, “Imperfect competition, compensating differentials, and rent sharing in the US labor market,” *American Economic Review*, 2022, *112* (1), 169–212.
- Lawless, Martina**, “Deconstructing Gravity: Trade Costs and Extensive and Intensive Margins,” *Canadian Journal of Economics*, 2010, *43* (4), 1149–1172.
- MacKenzie, Gaelan**, “Trade and Market Power in Product and Labor Markets,” 2019. University of Toronto.
- Naidu, Suresh, Eric A Posner, and E Glen Weyl**, “Antitrust Remedies for Labor Market Power,” *Harvard Law Review*, 2018, *132*, 536–601.
- Ruhl, Kim J and Jonathan L Willis**, “New exporter dynamics,” *International Economic*

*Review*, 2017, 58 (3), 703–726.

**Sedláček, Petr**, “Creative destruction and uncertainty,” *Journal of the European Economic Association*, 2020, 18 (4), 1814–1843.

**Simonovska, Ina and Michael E Waugh**, “Trade models, trade elasticities, and the gains from trade,” Technical Report, National Bureau of Economic Research 2014.

**Spencer, Adam Hal**, “Policy effects of international taxation on firm dynamics and capital structure,” *The Review of Economic Studies*, 2022, 89 (4), 2149–2200.

– , “Firing multinational CEOs,” *Economics Letters*, 2024, 239, 111766.

– , “Firm Dynamics and Multinational Activity,” in “Elgar Encyclopedia of International Trade,” Northampton, Mass.: Edward Elgar, Forthcoming.

**Tauchen, George**, “Finite state markov-chain approximations to univariate and vector autoregressions,” *Economics letters*, 1986, 20 (2), 177–181.

**Webber, Douglas A**, “Firm market power and the earnings distribution,” *Labour Economics*, 2015, 35, 123–134.

**Yeh, Chen, Claudia Macaluso, and Brad Hershbein**, “Monopsony in the US labor market,” *American Economic Review*, 2022, 112 (7), 2099–2138.

**Zlate, Andrei**, “Offshore production and business cycle dynamics with heterogeneous firms,” *Journal of International Economics*, 2016, 100, 34–49.