Trade liberalization and inequality: a dynamic model with firm and worker heterogeneity*

Matthieu Bellon†

January 2018

Abstract

Trade liberalizations are associated with higher wage inequality, but nearly the entire related literature is silent about the transition process. I address these limitations by developing a micro-founded model highligthing the reallocation dynamics between heterogeneous firms and workers. On the transition path following a trade reform, expanding high-paying exporters benefiting from new opportunities abroad increase wages to recruit better workers faster. Increased competition leads domestic firms' workers to accept wage cuts to delay their employers' exit and keep their job. I provide micro-empirical evidence supporting the main novel mechanisms. Results from the calibrated model suggest an overshooting response of inequality.

What are the dynamic effects of a trade liberalization reform on the wage distribution within a country? The vast majority of studies have focused on comparative statics and comparative steady states and shown that trade liberalizations are associated with rises in inequality. Is the response of inequality over time monotonic? If it is, how fast does it unfold? If it is non-monotonic, what is the nature, magnitude and length of the transitory effect? By nature, comparative statics and comparative steady states approaches are not suited to address these questions directly. I study the dynamic effect of trade on inequality by developing and calibrating a micro-

*I am grateful to Donal Davis, Jonathan Vogel, Amit Khandelwal and David Weinstein for their advice and guidance, seminar and conference participants at Columbia, NYU, Penn State, the University of Oregon USC, Dornsife, and the ETSG for comments and suggestions. I also would like to thank the Program for Economic Research and the Jerome A. Chazen Institute of International Business for financial support. This work is also supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'avenir" program (reference : ANR-10-EQPX-17 - Centre d'accès sécurisé aux données - CASD) . All remaining errors are mine.

†Mail: IMF, 700 19th Street NW, Washington, DC 20431. Email: bellon.matthieu@gmail.com

founded model with an explicit dynamic reallocation process.

Preliminary evidence from an event study suggests that the response of inequality could be hump-shaped.[1] I follow the evolution of inequality in response to liberalization reforms in 37 countries. Consistent with previous work,[2] figure 1 shows that inequality rises after countries open to trade. The increase is gradual over the first seven years. From this point, inequality stops increasing and the data suggest that this increase gets partially undone. By contrast, over the same period the ratio of trade to GDP rises steadily after the date of the liberalization. Country panel studies of this type obviously cannot identify the causal effects of trade.

The heart of my analysis is, therefore, structural. Results from a micro-founded trade model calibrated on labor micro-data also predicts inequality overshooting.

The dynamic general equilibrium model of trade developed in this paper encompasses both worker and firm heterogeneity. Heterogeneity on the two sides of the labor market fits squarely with key findings of the empirical literature that emphasize the dominant role of intra occupation-sector inequality, and positive assortative matching between heterogeneous firms and workers.[3] The product market, trade and firm heterogeneity are based on the workhorse model of Melitz (2003),[25] henceforth M03. Firm heterogeneity is revealed at birth and permanent. Worker heterogeneity follows Helpman et al. (2010),[20] henceforth HIR10. Worker ability is assumed to be unknown until a worker gets a job interview. The worker ability revealed at interview is specific to the interviewing firm, namely match specific.

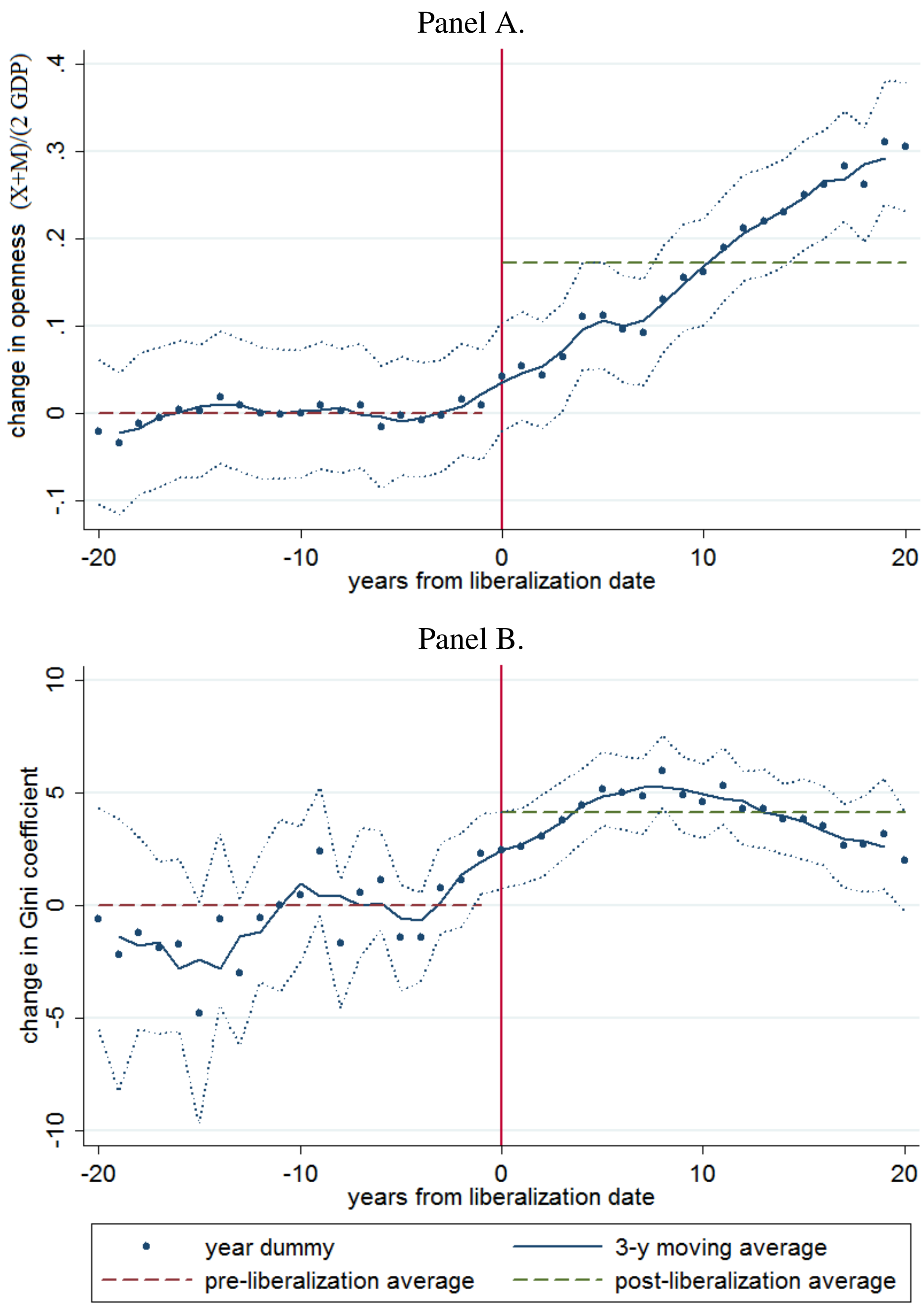
The dynamics of the model are generated by labor search frictions and costly recruiting costs. The model builds on the labor market with directed search developed in Kaas and Kircher (2015):[23] unemployed workers search for at most one vacancy type, and their probability of getting an interview depends on the ratio of searchers over vacancies. Once a worker is interviewed, hiring is decided depending on the revealed match productivity. Because of complementarity in the production

[1] This exercise is motivated by country case studies showing rich dynamic responses. Frias et al. (2009)[15] for Mexico and Helpman et al. (2012)[19] for Brazil both document a hump-shaped response of inequality following trade-enhancing reforms.

[2] See Goldberg and Pavcnik (2007)[17] for a survey of the effects of liberalization reforms on inequality in emerging countries.

[3] See Card et al. (2013)[6] and Helpman et al. (2012)[19] as discussed in the literature review.

Figure 1: Average evolution of openness and inequality after liberalizations.



Sources: WDI and WIID. The blue dots correspond to the estimates of time effects in a simple regression of the dependent variables on time dummies. The plain blue line is a 3-year moving average of these estimates. The dotted lines correspond to 95% confidence intervals of estimates. More details are provided in appendix.

function between the number of workers and worker ability, firms have incentives to screen workers. Because of additional complementarity with firm productivity, more productive firms are larger and screen more intensively.

In equilibrium, I obtain that job-searchers arbitrage between vacancy types that differ in matching rate, screening rates and expected payments. As in Kaas and Kircher (2015), high-wage jobs attract more workers and are harder to get; therefore, firms that want to hire faster do so by offering higher wages. My model adds to Kaas and Kircher (2015) homogeneous worker framework: firms that screen more intensively also need to increase wage offers in order to compensate job-searchers for the higher probability of being screened out. Better workers are paid more.

Firms do not leap to their optimal state; instead, they grow gradually in order to save on vacancy posting costs that are assumed to be convex. Moreover, the further away from its optimal size a firm is, the larger the opportunity costs and the faster it grows. By assumption firms start small with an unselected set of workers. They grow production capacities by screening the least able of their workers and by replacing them with more and better workers. Firms decide to export only if they are productive enough and only when they are big enough to cover the fixed costs of export. Even when aggregates are constant and the economy is in steady states, the model features overlapping generations of gradually growing firms.

I extend steady state results of the literature with static firms by showing that welfare and inequality in the dynamic model are larger under trade than under autarky. Moreover, I prove equivalence results with canonical models. In particular, I derive conditions under which the elasticity of steady state consumption with respect to trade openness is the same as in M03. I also show that the relationship between the steady state dispersion of wages and the level of trade costs has the same inverted U-shape as in HIR10. I also prove that steady state unemployment increases and employment in the tradable sector decreases with openness.

The explicit firm dynamics in the model allows me to derive results for individuals and firms on the transition path from one steady state to another. A once and for all reduction in trade costs, provides new opportunity for current or future exporters while greater competition reduces revenues for domestic firms.

Among the firms born before the reform, the larger and more productive ones immediately export, they increase employment and screen more intensively at a faster rate than what they would have without the reform. They do so because they are now further away from their new optimal size. They also offer higher wages than they would have, because that speeds adjustments, the hiring process, and allows them to select better workers. I also show that at some point in the transition, offers can be so large that some firms pay wages that exceed what the firms with the same age and characteristics will offer in the future steady state. I also prove that the wage dispersion within these firms temporarily increases during the transition while new workers are hired at a premium.

By contrast, the smaller and less productive firms born before the reform that will not export grow more slowly, shrink or exit. Because of slower growth and looser screening, domestic firms that still hire offer lower wages. Furthermore, there are some firms that need to cut costs to remain profitable and survive: I assume that they negotiate wage cuts. Workers only accept cuts when these allow them to prevent the firm exit and to keep their job with a wage above their outside option. Some of least productive firms that are able to survive for a while are gradually replaced by more productive firms.

I provide new micro empirical evidence of the model main novel mechanisms and use estimation results in a calibration exercise to compute the evolution of aggregate variables. I use a matched employer-employee data from the French manufacturing sector over 1995-2007. I argue that French labor regulation being average by standard measures, the French data is reasonably appropriate for the study. I find that worker separation rates increase with firm growth. This is consistent with the prediction that firms grow production capacity by increasing the number *and* average ability of workers, including by screening out the least able of their existing workers. I also use an IV estimation approach to show that wage offers increase with firm growth as in the model, with an estimated elasticity of 10%.

Numerical results from the calibrated model show that, in response to the opening of the economy, wage dispersion in the tradable sector overshoots its new long run level. The model predicts that inequality peaks four years after the reform and that about one fourth gets undone in the following ten years. Variation in inequality is mostly driven by the dispersion of firm average wages. At the top of the wage distribution, temporary wage premia are offered by high-paying expanding firms to speed the adjustment process. In the long run, employees hired at a premium retire and are replaced by workers without such premia. At the bottom of the wage distribution, the least productive firms cut wages. In the long run, these firms exit and are replaced by better firms that pay more on average. Within firm inequality also contributes to a temporary increase in inequality as the wage dispersion increases at expanding firms, but its contribution to total inequality is small.

This paper contributes to a growing literature that focuses on the interplay between labor frictions, firm growth and trade. Despite having explicit endogenous firm dynamics, the work in Fajgelbaum (2013),[13] Felbermayr et al. (2014),[14] or Cosar et al. (2016) only compare steady states. In addition to considering the equilibrium out of steady state, my model improves on Kaas and Kircher (2015)[23] by adding worker heterogeneity and embedding the model in an economy with differentiated products and the possibility to trade.

This paper that focuses on within sector-skill group inequality complements the out-of-steady state models that focus on gradual reallocation of workers across sectors (Cosar (2010)[8] and Dix-Carneiro (2010)[11]) or on gradual skills acquisition (as in Emami-Namini and Lopez (2013)[27] and Danziger (2014),[9] two papers predicting the overshooting of the skill premium). While sectoral differences can be important, previous empirical work has shown that most of the reallocation flows and inequality variations occur within sectors (Wacziarg and Wallack (2004),[29] Helpman et al. (2012)[19]). In addition, within skill group inequality is highlighted as a key component of overall inequality variation by the empirical literature.[4]

To this date, only one paper features changes in wage dispersion along the transition path in a model of overlapping generations of firms. Helpman and Itskhoki (2015)[?] develop a trade model with search frictions and bargaining where dynamics are driven by exogenous quits and firm exits, and where changes in wage

[4] Card et al. (2013)[6] for Germany: "Increasing workplace heterogeneity and rising assortativeness between high-wage workers and high-wage firms likewise explain over 60% of the growth in inequality across occupations and industries." Helpman et al. (2012)[19] for Brazil: "The within sector-occupation component of wage inequality accounts for over two thirds of both the level and growth of wage inequality in Brazil between 1986 and 1995." Frias et al. (2009)[15] find in Mexican data that wage changes result from changes in plant premia rather than changes in skill composition.

inequality are driven by wage cuts at the least productive declining firms. A similar channel operates in my model. However, two features set the models apart. First, the absence of worker heterogeneity in their model implies that there is no wage dispersion in steady states. Second, all expanding firms in their framework leap instantaneously to their long run optimal states and pay wages that are constant across firms and over time. Thus, their model rules out any wage dispersion across firms and any wage changes at the top of the firm productivity distribution.

The rest of the paper is organized as follows. Section 1 lays out the mechanisms of the labor market, the trade-offs faced by job-searchers and the optimal behavior of firms. Section 2 is dedicate to general equilibrium steady state properties and equivalence results with the literature are provided. In section 3, I derive some properties for the transition path following a once and for all reduction in trade costs. Section 4 presents micro-empirical evidence, the calibration, and discusses the numerical results of the calibrated model. Section 5 concludes.

1 Recruitment and worker selection by growing firms

The model is set up in discrete time. I consider two symmetric countries, domestic (D) and foreign (F), in which there are two sectors. Because of symmetry, I choose to focus on the home country only. There is no aggregate shocks.

One sector uses labor to produce and sell a homogeneous good to the other sector under perfect competition. This good is non-tradable and can be thought of as services. Workers in the service sector are hired in a perfectly competitive labor market and the competitive wage is the numéraire.[5]

In the other sector, an infinite number of heterogeneous firms produce and sell varieties of the final good. Firms in this sector are searching for workers with specific ability in a labor market that is characterized by frictions.

Workers and firm owners have access to an insurance market that is complete except that it cannot insure against a trade reform. The latter is assumed to be unexpected by all agents and is studied later separately. Nevertheless, insurance implies that probability-weighted sums of payment streams are sufficient for individuals to

[5] It is equal to productivity in services and assumed to be constant over time.

evaluate different decisions. In period t, individuals discount future income flows in $t' > t$ using the interest rate $R_{t,t'}$. In steady states, the latter is denoted by R_S.

1.1 Labor supply

Workers can either be unemployed or employed, but only the currently unemployed workers can look for job vacancies. I assume that firms are large relative to workers in the sense that each firm employs a continuum of workers and that the law of large numbers applies for all probabilistic events.

1.1.1 Job searchers arbitrage across different vacancy types

Workers search for vacancies that differ along three dimensions. First, the probability of getting an interview, or equivalently the vacancy-finding rate, varies depending on the number of other applicants per vacancy. Second, not all interviews are successful as firms screen interviewees based on their ability. Third, job vacancies differ in stochastic streams of wage payments as summarized by job values.

The labor market is segmented by vacancy type, where a type is defined as a unique combination of a vacancy-finding rate, a screening rate and a job value. Vacancy types are publicly announced by firms.

Job-seekers can only search and apply to one type of vacancy per period. For each type, the vacancy meeting rate, m, is a standard function of the ratio of searchers to vacancies, Λ, such that $m(\Lambda) = \frac{c_m}{(1-\lambda)}\Lambda^{1-\lambda}$ with $0 < \lambda < 1$. Thus, the vacancy-finding rate is $\frac{m(\Lambda)}{\Lambda}$.

Worker ability is match-specific[6] and follows HIR10. Unemployed workers are ex-ante identical. Once they meet a firm, they draw an ability level α. Ability draws are assumed to be independently and identically distributed across firm-worker matches and unknown to all until workers take a test.

Firms posting vacancies announce how selective their job openings are. They specify the ability cutoff α_c below which interviewees will be screened out. Specifically, the interview is a test revealing whether a worker's ability is below α_c. The

[6]When a worker separates from a match and goes back to the unemployment pool, her ability at her past match will be completely unrelated to her ability draws at future interviews. This rules out any self selection into different vacancy types and keeps the model tractable.

distribution of ability among workers follows a Pareto distribution with shape parameter $\kappa > 1$ and lower bound α_{min}. Hence, the probability of being selected at an interview is equal to $(\alpha_{min}/\alpha_c)^{-\kappa}$ and ex-ante identical for all applicants.

Unemployed workers can either be inactive or incur a search cost c_u with the possibility of getting hired. Job searchers that get hired start working immediately. All workers without a job at the end of the search period (whether there was any search or not) get some government benefit b_u. The value $V_{u,t}$ of being unemployed at t and applying to a vacancy of type (Λ, α_c, V_j) where V_j is the job value is:

$$V_{u,t}(\Lambda, \alpha_c, V_j) = \frac{m(\Lambda_t)}{\Lambda_t}\left(\frac{\alpha_{min}}{\alpha_{c,t}}\right)^{\kappa} V_{j,t} - c_u + \left(1 - \frac{m(\Lambda_t)}{\Lambda_t}\left(\frac{\alpha_{min}}{\alpha_{c,t}}\right)^{\kappa}\right)\left[b_u + \frac{V_{u,t+1}}{R_{t,t+1}}\right] \quad (1)$$

while the value of being inactive (not applying) is $V_{u,t} = b_u + V_{u,t+1}/R_{t,t+1}$.

Because workers are free to direct their search towards any type of vacancy, the value of being unemployed is equalized across all types and activity status in equilibrium: $V_{u,t}$ is independent of the vacancy type (Λ, α_c, V_j).

1.1.2 Job values depend on wages and job continuation probabilities

In every sector, employed workers incur a constant dis-utility c_j, with $j = e$ in the differentiated sector, and $j = r$ in services. Employment values depend on this dis-utility parameter, the wage w_t, the job termination rate η_{t+1} and the option value of still being employed next period:

$$V_{j,t} = w_t - c_j + \frac{1}{R_{t,t+1}}\left[(1-\eta_{t+1})V_{j,t+1} + \eta_{t+1}V_{u,t+1}\right] \quad (2)$$

There are several ways for employed workers to lose their job. At the minimum, jobs are destroyed when firms and workers are hit by exogenous exit and separation shock respectively. In addition, a firm can fire and screen workers out based on ability. Thus, job continuation rates $(1-\eta_{t+1})$ are at most equal to some probability $(1-\eta_S)$ and can be greater depending on firm screening and lay-off policies.

Job-searchers choose freely to look for any job, either in the service or the differentiated good sector. This model focuses on the differentiated sector: I assume that workers find jobs and get hired in services without friction and delay.[7] Conse-

[7] This can be achieved by assuming that λ goes to infinity in the service sector.

quently, service sector vacancies are such that $m(\Lambda)/\Lambda = 1$, $\alpha_c = \alpha_{min}$, and workers can keep working in service jobs forever. In services, the value function simplifies to $V_{r,t} = 1 - c_r + V_{r,t+1}/R_{t,t+1}$.

The unemployed arbitrage between getting a service job immediately or looking for a job in the differentiated sector. I obtain in equilibrium:

$$V_{u,t} = V_{r,t}, \quad \text{and} \quad b_u + c_r = 1 \tag{3}$$

1.1.3 Labor supply equations link wages to market tightness and screening

In the differentiated sector with searching and screening, there is an infinity of wage schedules that allow a firm to deliver a value $V_{e,t}$ at the same cost for the same firing and screening schedule because wage payments can indifferently be made sooner or later.[8] Indeed, wages are defined in equation (2) as a implicit function of the current and future the job values, and the future separation and interest rates: $w(V_{e,t}, \eta_{t+1}, V_{e,t+1}, R_{t,t+1})$.

To resolve the indeterminacy, I assume that the above wage function is independent of future variables,[9] and therefore only depends on the current job value.[10] I show in appendix that the only wage function satisfying this assumption is:

$$w_t \quad = \quad 1 - c_r + c_e + [V_{e,t} - V_{r,t}]/c_w, \quad \text{where} \quad 1/c_w \equiv 1 - (1-\eta_S)/R_S \tag{4}$$

Equipped with the wage function (4), the initial wage of new hires can be related to the vacancy and labor market characteristics. Job-searchers' arbitrage condition across vacancy types (3) implies that hard-to-get jobs provide a higher initial wage

[8] The original directed search model of Kaas and Kircher (2015) [23] has the same indeterminacy. This can be seen in equation (2) where for given $\{\eta_s\}_{s>t}$, a firm can deliver a job value $V_{e,t}$ with any current wage w_t as long as it commits to an appropriate future job value $V_{e,t+1}$.

[9] In an environment with computation costs or with some costs of acquiring information about future changes, individuals would favor such restriction.

[10] This assumption generalizes the assumption in Kaas and Kircher (2015) [23] to a dynamic environment with selection as it nests their assumption that wages at firms in stationary equilibria are constant over time. The authors note that their assumption corresponds to the limit case of risk-neutral firms and risk-averse workers as risk aversion vanishes.

(or job value):

$$w_t = \underbrace{1 - c_r + c_e}_{\equiv w_o} + \underbrace{\frac{\Lambda}{m(\Lambda)}\left(\frac{\alpha_{min}}{\alpha_c}\right)^{-\kappa}\frac{c_u}{c_w}}_{\text{"the wage premium"}} \tag{5}$$

Equation (5) expresses that workers must at least be paid an amount w_o that combines their outside option and a compensation for the dis-utility of working in the differentiated sector. Furthermore, firms can only attract workers if they compensate them for search-related costs with a "wage premium". This second term increases with the average time to get an interview and the screening rejection rate. It means that more productive workers that went through a stricter screening process are rewarded with higher wages.

After the initial wage payment, wages continue to vary with the labor market environment and firm labor management. Firms are bound to deliver wage payments that are consistent with the commitments they made at the time t of hiring. I show that:

$$w_{t'+1} - w_o = \frac{R_{t',t'+1}(1-\eta_S)}{R_S(1-\eta_{t+1})}(w_{t'} - w_o), \qquad t' > t \tag{6}$$

Equation (6) means that the workers in a cohort hired in period t that are kept after lay-offs (when $\eta_{t'+1} > \eta_S$) get a wage raise ($w_{t'+1} > w_{t'}$). Such raises ensure that firms compensate for the losses associated with lay-offs, thereby delivering on their commitment at hiring to provide a job value $V_{e,t}$ in expectation.[11]

In summary, wage equations (5) and (6) characterize initial wage offers and subsequent wage changes as a function of firm decisions. As in Kaas and Kircher (2015),[23] jobs that are harder to get because they attract more candidates pay higher wages. In addition, the selection process that is absent in their paper entails that firms that screen more need to compensate their workers with higher wages.

1.2 Labor demand and firm profits maximization problem

In this section, I drop references to periods and t subscripts for the sake of clarity.

[11] For the same reason, an increase in the interest rate reduces future job values and is compensated by wage raises. As an increase in the interest rate is the result of an increase in borrowing demand in a growing economy, it implies that all wages increase when the economy increases.

1.2.1 Overlapping generations of firms under monopolistic competition

Entrepreneurs may create firms by paying a sunk cost f_e. Once the cost is paid, a productivity x is drawn from a Pareto distribution with c.d.f $F(x)$, shape parameter $\theta > 1$ and lower bound x_{min}. Firm productivity x remains constant until exit. Upon entry, firms get a small initial number of workers, $l_0(x, \mathscr{A})$, whose ability is Pareto-distributed above the cutoff $\alpha_{c,0}(x, \mathscr{A})$, where $\mathscr{A}$ represents aggregate variables.

All firms are hit by an exogenous exit shock with probability δ_S before any other decisions are made. In addition, new firms choose $\delta \in \{\delta_S, 1\}$ capturing the decision about whether to exit ($\delta = 1$) or to enter the market and produce ($\delta = \delta_S$) if they survive the exogenous shock. After a successful entry, complete insurance markets imply that firms never choose to exit except at the trade reform announcement. Because I assume that the trade reform is unexpected and cannot be insured against, exit in that period can be a rational decision as studied in section 2.1.1.

The production q of each firm takes place at the end of periods and depends on the firm specific productivity x, the number of employed workers l', and their average ability $\overline{\alpha}'$:

$$q = x\overline{\alpha}' l'^{\rho}, \qquad 0 < \rho < 1 \tag{7}$$

Primes are used to distinguish values at the end of periods from the beginning of periods. The technology corresponding to the production function features complementarities between the three variables.[12]

Talented workers need the appropriate support technology and environment to fully reach their potential productivity. Providing this is costly, requiring firms to pay a support cost $c_s \overline{\alpha}'^{\psi}/\psi$ every period that is increasing in workers' ability ($0 < \psi$).[13] Production also requires a fixed cost f_d per period.

Demand $\mathscr{C}^M$ in each country $M \in \{H, F\}$ over the continuum of varieties within the sector is standard and takes the constant elasticity of substitution form with

[12] The function, also used in HIR10, is log supermodular: an increase in any of the three variables yields a greater increase in output when the other variables are larger.

[13] Even if this cost function has the exact functional form as the "screening cost" in HIR10, their interpretations differ. HIR10's "screening cost" is described as the per-period cost of acquiring the technology to screen job-candidates. However, this "screening cost" interpretation is at odds with the fact that firms pay this cost every period, including when they are not hiring anybody. With explicit firm dynamics, the "support cost" interpretation is preferable: it is an expense that is related to the production occurring in every period rather than to the occasional hiring and screening process.

parameter $\sigma > 1$. I introduce a market competition index $\phi^M \equiv (\mathscr{C}^M)^{1/\sigma}\mathscr{P}^M$ combining aggregate demand $\mathscr{C}$ and the associated price index $\mathscr{P}$ which depends on the prices p of individual varieties.[14]

When trade is possible, firms have the option to pay a fixed costs f_X and a variable "iceberg" cost τ to sell abroad. The productions intended for the domestic and foreign markets are respectively represented by q^D and q^F with $q = q^D + q^F$. Firm total revenues are the sum of those on each market which amount respectively to $r^D = \left(q^D\right)^{\frac{\sigma-1}{\sigma}}\phi^D$ and $r^F = \left(q^F/\tau\right)^{\frac{\sigma-1}{\sigma}}\phi^F$.

1.2.2 Active firms manage their workforce by recruiting, screening and firing

I make two assumptions about screening. First, I assume that firms use the same screening cutoff α'_c for new hires and incumbent workers alike. This implies that firm average ability is $\overline{\alpha}' = \alpha'_c \frac{\kappa}{\kappa-1}$. Equivalently, I could impose prohibitive additional costs or productivity losses for keeping workers with ability levels that are too low compared with the firm average ability.[15]

The second assumption is that firms cannot decrease the screening cutoff unless macro conditions unexpectedly deteriorates, in which case they may decrease it to the firm stationary cutoff level. I show in the next subsection that this assumption is not binding in steady state environments: all steady state results hold without the assumption. However, the assumption is needed to formulate the firm problem in recursive form. The trade reform studied in section 2.3 is the only unexpected change in macro conditions considered. At the reform, the assumption becomes binding for some firms but it allows me to keep the problem tractable. I discuss the assumption in section 2.1.1 in more details. Formally, it corresponds to $\alpha'_{c,\infty} \geq \alpha'_c \geq min(\alpha_c, \alpha'_{c,\infty})$ where $\alpha'_{c,\infty}(x;\mathscr{A})$ denotes the stationary cutoff level.

In addition to screening, firms can also fire a fraction of workers of all ability levels equally by choosing a separation rate s above the exogenous quit rate s_S. As such, firing differs from screening because it leaves average ability unchanged and avoids a raise in support technology costs. Firing at all ability levels and screening are cumulative, and once the survival probability $(1-\delta)$ of employing firms is also

[14] As is well known, the demand function for a given variety can be expressed as $(p^M/\phi^M)^{-\sigma}$.

[15] Specifically, the assumption can be micro-founded by adding a large enough penalty for firms having workers with ability below a threshold equal to $\frac{\kappa-1}{\kappa}$ times their workforce average ability.

accounted for, the job continuation rate is expressed as:

$$(1-\eta) = (1-\delta)(1-s)\left(\frac{\alpha_c'}{\alpha_c}\right)^{-\kappa}, \qquad s \geq s_S,\;\; \delta \geq \delta_S \tag{8}$$

Searching for workers and recruitment can be achieved by posting vacancies. It is a costly process and vacancy costs C are assumed to depend on the number V of vacancies and the number $\check{l}$ of workers available to help in the recruiting process. Because all separations occur first, these workers are the workers l that remain from last period after lay-offs and screening.

$$C(V,\check{l}) \;\; = \;\; \frac{c_c}{1+\nu}\left(\frac{V}{\check{l}}\right)^{\nu} V, \qquad c_c > 0,\;\; \nu > 0,\;\; \check{l} \equiv (1-s)\left(\frac{\alpha_c'}{\alpha_c}\right)^{-\kappa} l \tag{9}$$

This functional form has been used extensively in previous theoretical labor work.[16] Vacancy costs are increasing in the number of vacancies as well as in the vacancy rate $V/\check{l}$. Costs increase convexly with the number of vacancies.[17]

The number Δ of new recruits depends on the number of vacancies, the searcher-vacancy ratio through the matching function, and the intensity of interview screening. A number V of vacancies result in $m(\Lambda)V$ interviews and the selection of $\Delta = m(\Lambda)V(\alpha_c'/\alpha_{min})^{-\kappa}$ new hires. More recruits can be obtained either with more vacancy openings or with higher wage offers attracting more workers per vacancy as stated in (5). Using the matching function introduced in the 1.1.1, I obtain:

$$\Delta \;\; = \;\; c_m \frac{\Lambda^{1-\lambda}}{1-\lambda} V \left(\frac{\alpha_{min}}{\alpha_c'}\right)^{\kappa} \tag{10}$$

$$w \;\; = \;\; w_o + \frac{(1-\lambda)\Lambda^{\lambda}}{c_m}\left(\frac{\alpha_c'}{\alpha_{min}}\right)^{\kappa}\frac{c_u}{c_w} \tag{11}$$

The size l of a firm workforce is the result of recruiting, screening and firing decisions:

$$l' = \check{l} + \Delta = (1-s)\left(\alpha_c/\alpha_c'\right)^{\kappa} l + \Delta \tag{12}$$

Total wage payments to new hires and incumbent workers depend on firms'

[16] Merz and Yashiv (2007)[26] and in Kaas and Kircher (2015)[23] use a similar function.

[17] These features are supported by many empirical studies. Blatter et al. (2012)[5] and Dube et al. (2010)[12] all document increasing marginal recruitment costs. Blatter et al. (2012)[5] and Barron et al. (1997)[4] also find that larger firms have higher hiring rate costs. See Manning (2011)[24] for a survey of the literature.

current and past commitments. There are two wage bill components. First, all firms have to pay $w_o l'$, the outside option for all workers. Second, firms need to pay each worker cohort their promised premium. Let B be the beginning-of-period total wage premium commitments $(w - w_o)$ accumulated by a firm. Every period, past wage premium commitments vary according to wage equation (6), decrease with the number of separations and increase with new commitments:

$$B' = \underbrace{B\frac{R'(1-\eta_S)}{R(1-\eta)}}_{\text{wage variation (6)}} \underbrace{(1-s)\left(\frac{\alpha_c}{\alpha_c'}\right)^{\kappa}}_{\text{separations}} + \underbrace{\Delta(w-w_o)}_{\text{new commitments}} = B\frac{R'}{R}(1-s_S) + \Delta.(w-w_o) \tag{13}$$

I substitute the continuation rate $(1-\eta)$ using (8) in the second equality. The simplified expression shows that the wage premium bill does not decrease with firing or screening. The wage premia that used to be paid to actively laid-off workers is exactly equal to the wage raises of retained incumbents: B is unaffected. Note that firing and screening nevertheless imply some wage costs reduction via a decrease in the outside option component.

1.2.3 The firm problem

Per-period gross profits π are equal to the revenues net of the vacancy and support technology costs, and net of the wage bill corresponding to the payments of workers' outside option plus the sum of promised wage premia B'.

$$\pi = r - C - \frac{c_s}{\psi}\overline{\alpha}'^{\psi} - w_o l' - B' \tag{14}$$

Net profits Π are obtained from subtracting the fixed costs from gross profit: $\Pi = \pi - f_d - I_X f_X$.

Active firms maximize G, the expected value of their current and future net profits taking aggregate demand and prices as given. To do so, they make recruiting decisions with respect to the number of vacancies to open, the overall ability screening cutoff, the number of new hires and their wage, and the level of firing at all ability levels. They decide on production levels and whether to export, in which case the dummy variable I_X is equal to one.

$$\begin{aligned} G(\alpha_c, l, B, x; \mathscr{A}) \quad &= \quad (15) \\ \max_{\substack{V\geq 0,\ \Delta\geq 0,\\ \Lambda\geq 0,\ q^D, q^F\geq 0,\\ s\geq s_S,\ \alpha_\infty\geq\alpha_c',\\ \alpha_c'\geq min(\alpha_c,\alpha_{c,\infty}'),\\ w\geq w_o,\ I_X\in\{0,1\}}} (1-\delta_S)&\left\{ r(q^D, q^F; \mathscr{A}) - C(V, \check{l}) - \frac{c_s}{\psi}\overline{\alpha}'^{\psi}_c - w_o l' - B' - f_d - I_X f_X + \frac{1}{R'} G\left(\alpha_c', l', B', x; \mathscr{A}'\right) \right\} \\ &\text{subject to equations (7), (9), (10), (11), (12) and (13)} \end{aligned}$$

The aggregate variables $\mathscr{A}_t = \{\phi_{t'}, R_{t',t'+1}\}_{t'\geq t}$ determining firms decisions can be complex in non stationary environments because they encompass current and future aggregates. However, intra-temporal decisions $(q, I_X, V, \Lambda, w, \Delta)$ are only affected by current aggregate variables and firm beginning-of-period variables.

1.2.4 Intra-temporal decisions: labor management, production and export

Every period, exporters allocate their sales optimally. They equate the marginal revenues across markets given their production capacity q. This leads exporting firms to allocate a share $1/\Upsilon$ of their production and revenues to the domestic market, with $\Upsilon \equiv 1 + (1/\tau)^{\sigma-1}\left(\phi^F/\phi^D\right)^{\sigma}$. Export quantities and revenues are respectively $q^F = q(\Upsilon-1)/\Upsilon$ and $r^F = r(\Upsilon-1)/\Upsilon$. Total revenues for exporters and domestic-only firms take the form:

$$r(q, I_X; \phi^D, \Upsilon) = q^{\frac{\sigma-1}{\sigma}} \phi^D \left[1 + (\Upsilon-1) I_X\right]^{\frac{1}{\sigma}}, \tag{16}$$

Not all firms find it profitable to exports. Only the firms that sell enough to cover the fixed export costs decide to export.[18]

$$I_X = \begin{cases} 1 & \text{if } r(q,1;\phi,\Upsilon) - f_X \geq r(q,0;\phi,\Upsilon) \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

As in other recent dynamic models,[19] exporting firms are those productive and old enough to have accumulated enough production capacities.

[18]From now on the superscript ^{D} is dropped because countries are assumed to be symmetric.

[19]Holzner and Larch (2011),[21] Felbermayr et al. (2014)[14] and Fajgelbaum (2013)[13] to name a few. Both in their models as well as in the present setup, there is not a single export cutoff as in M03 but an export cutoff function that depends on firm accumulated production capacities.

Given beginning-of-period l and α_c, and having chosen Δ and α'_c, firms decide on how to optimally recruit Δ workers with an ability level α'_c or greater. Firms trade off between posting more vacancies and attracting more workers per vacancy with higher wages. Optimality conditions yield

$$
\begin{aligned}
\Lambda(\Delta,\alpha'_c,\check{l}) &= c_\Lambda \left(\Delta/\check{l}\right)^{\frac{\gamma(\gamma-1)}{\gamma-1+\xi}} \left(\alpha'_c/\alpha_{min}\right)^{\frac{\kappa\gamma(\gamma-1)}{\gamma-1+\xi}} && (18)\\
V(\Delta,\alpha'_c,\check{l}) &= c_V\, \check{l} \left(\Delta/\check{l}\right)^{\frac{\gamma\xi}{\gamma-1+\xi}} \left(\alpha'_c/\alpha_{min}\right)^{\frac{\kappa\gamma\xi}{\gamma-1+\xi}} && (19)
\end{aligned}
$$

with $\gamma = (1+\nu)/(1+\nu(1-\lambda)) > 1$ and $\xi = \lambda/(1+\nu(1-\lambda)) \in (0,1)$.[20] Faster growing firms (characterized by a high recruitment rate $\Delta/\check{l}$) and more selective firms choose both to have a higher vacancy rate ($V/\check{l}$) and to attract more candidates per vacancy (Λ).

Faster employment growth and improvements in workforce ability imply higher adjustment costs, both for the firm and for the rest of society. For recruiting firms, these costs are denoted by A and consist of the vacancy posting costs plus the present value of wage premia commitments: $A \equiv C + c_w\Delta(w - w_o)$. For the rest of the economy, greater firm adjustments also translate into longer queues per vacancy. Raising wage offers above the workers' outside option allows firms to grow faster and to be more selective but also generates unemployment by attracting more candidates that won't get a job. For every recruiting firm and the corresponding vacancy type, the number u of searching workers is the product of vacancies and the ratio of searchers to vacancy. The number u' of unsuccessful job searchers is $u-\Delta$, after the number of matches are subtracted. Formally, I show that:[21]

$$
\begin{aligned}
w(\Delta,\alpha'_c,\check{l}) &= w_o + \frac{(1-\xi)}{c_w}\frac{A}{\Delta}, \qquad \text{with } A(\Delta,\alpha'_c,\check{l}) = \frac{c_A}{\gamma}\Delta\left(\frac{\Delta}{\check{l}}\right)^{\gamma-1}\left(\frac{\alpha'_c}{\alpha_{min}}\right)^{\kappa\gamma} && (20)\\
u'(\Delta,\alpha'_c,\check{l}) &= \frac{1-\xi}{c_u}A - \Delta && (21)
\end{aligned}
$$

Adjustment costs per new hire (A/Δ), wage offers, as well as the ratio of rejected candidates (u'/Δ) are increasing in the firm hiring rate ($\Delta/\check{l}$) and in the firm screening intensity. Rewriting equation (20) shows that a constant fraction ξ of firm

[20] Additionally, $c_\Lambda \equiv \left(\frac{c_C}{c_u}\frac{(1-\lambda)^{1+\nu}}{\lambda\ c_m^{\nu}}\right)^{\frac{1}{1+\nu(1-\lambda)}}$ and $c_V \equiv \frac{c_\Lambda^{\lambda-1}(1-\lambda)}{c_m}$. See the appendix for details.

[21] $c_A \equiv \left(\frac{1-\xi}{\gamma}\right)^{\gamma-1}\left(\frac{1-\xi}{\gamma-1+\xi}\right)^{\xi}\frac{c_u^{1-\xi}c_C^{\xi}}{c_m^{\gamma}}$ as shown in the appendix.

adjustment costs is posting costs, $C(\Delta, \alpha_c', \check{l}) = \xi A$, while the rest consists of wage premia. This fraction is increasing in the elasticity of the matching function (λ), and decreasing in the degree of convexity of the vacancy costs (ν). Unemployment is lower when the cost of searching (c_u) is higher as workers favor faster matching rates, and when matching is more efficient (as captured by c_m, and ξ). Conversely unemployment is higher when the cost of posting vacancy (c_C) is higher as firms opt for longer queues and higher vacancy filling rates.

1.2.5 The saddle path equilibrium of firm inter-temporal decisions

The following proposition and its corollaries characterize firm equilibrium paths and stationary states. The proofs are in appendix.

Proposition 1 (Firm saddle path equilibrium)
If $\gamma\kappa\rho \leq 1$ and $\psi > \Psi_1$,[22] the firm problem has a unique stationary equilibrium characterized by $\{\alpha_S(x,\phi,R,\Upsilon), l_S(x,\phi,R,\Upsilon), B_S(x,\phi,R,\Upsilon)\}$ and locally, around any firm-specific stationary point and assuming that initial conditions are such that $(\alpha_c - \alpha_S)(l - l_S) \geq 0$, there is a unique local saddle path equilibrium that leads to the firm stationary equilibrium.

Firm stationary states depend on productivity x and the aggregate environment.

Corollary 1 to proposition 1 (Firm stationary equilibrium)
In stationary states, all variables $\left(\alpha_S, l_S, q_S, I_{X,S}, r_S, \pi_S, \Delta_S, V_S, w_S, [u'/\Delta]_S, \Lambda_S, [V/\check{l}]_S\right)$ increase in x, ϕ and Υ, except for the separation and hiring rates $(s_S, [\Delta/\check{l}]_S)$ that are independent of (x,ϕ,R,Υ).

In their stationary state, all firms need to replace the constant fraction s_S of workers that exogenously quit: the hiring rate equals the quit rate ($\Delta_S/l_S = s_S$).

[22] The results hold for any (γ,ρ) when there is no screening (that is when ψ or κ goes to ∞, in which case $\alpha_c = \alpha_m$ always). In the no-screening case, I also prove in appendix that the saddle path is a global equilibrium. With screening, I cannot rule out that some initial conditions lead to diverging solutions. The first condition ($\gamma\kappa\rho \leq 1$) means that the returns to screening are high enough. Specifically, $\Psi_1 \equiv \max\left\{ \frac{\frac{\sigma-1}{\sigma}\frac{(1-\kappa\rho)^2}{\kappa\rho}\left[1-\frac{(1-\delta_S)}{R_S}\left(1-\frac{s_S}{\gamma}\right)\right]-\kappa\frac{(\gamma-1)^2}{\gamma}s_S}{\Omega_1}, \frac{\frac{\sigma-1}{\sigma}\left[\frac{1-\kappa\rho}{\kappa\rho}-\left(1-\frac{s_S}{\gamma}\right)s_S\right]\left[1-\frac{(1-\delta_S)}{R_S}\left(1-\frac{s_S}{\gamma}\right)\right]+\kappa\frac{\gamma-1}{\gamma}s_S(1-s_S)}{\Omega_1}, \frac{\sigma-1}{\sigma}(1-\kappa\rho)^2-\kappa^2\rho, \kappa\gamma+\frac{\frac{\sigma-1}{\sigma}}{1-\rho\frac{\sigma-1}{\sigma}} \right\}$ with $\Omega_1 \equiv \frac{(1-\kappa\rho)}{\kappa\rho}\left[1-\frac{(1-\delta_S)}{R_S}\left(1-\frac{s_S}{\gamma}\right)\right]-\frac{\gamma-1}{\gamma}s_S$ is the threshold derived in appendix.

Because of the triple complementarity between firm productivity, employment and workforce average ability, more productive firms decide to hire more workers and to screen them more intensively. Hence, by all measures (employment l, production q, revenues r, profits π), more productive firms are larger. Their larger sales are more likely to cover the fixed export costs and they are thus more likely to export. They also post more vacancies (V) and hire more workers (Δ). Because their screening rate is higher and the ratio of rejected candidates $([u'/\Delta])$ is higher, more productive firms need to have a higher vacancy posting rate $([V/\check{l}])$ and to attract more candidates per vacancy (Λ). The latter is achieved by offering higher wages.

All measures of firm size also increase with the easiness of competition (ϕ) and trade (Υ) that act like productivity boosters. Thus, firm variables that increase in x also increase with (ϕ, Υ) in the same way.

Firms that are not at their stationary state do not reach it in one period. Because of convex adjustment costs, it is more profitable to smooth adjustments over time and converge gradually to the stationary state. Using linearization around the stationary equilibrium (in appendix), I characterize gradual adjustments as follows.

Corollary 2 to proposition 1 (Firm variables on the saddle path)

(i) All firm decisions are independent of past wage premia commitments B.

(ii) For a given firm on its local saddle path equilibrium,[23] $[\Delta/\check{l}]$ *decreases as firm size measured by* $h \equiv \alpha_c^{\kappa} l$ *gradually grows to its stationary level from below. Meanwhile, the firing rate remains constant:* $s = s_S$. *Conversely,* $[\Delta/\check{l}]$ *increases and* s *decreases (weakly) when the firm shrinks to its stationary level from above.*

(iii) If the stricter conditions $\gamma\kappa\rho \leq \kappa\rho + (1-\kappa\rho)\rho\frac{\sigma-1}{\sigma}$ *and* $\psi > \Psi_2$ *apply,*[24] *then* (α_c, l, q, I_X, r) *rise and* $(w, \Lambda, [\Delta/\check{l}], [V/\check{l}], [u'/\Delta])$ *decrease with* h *on the saddle path.*

Because none of the firms' actions can affect the level of past wage commitments, firms do not take them into account when managing their workforce and deciding on a production strategy.

[23] In the no-screening case when either ψ or κ goes to ∞, the results are valid globally. See 22.

[24] Specifically, $\Psi_2 \equiv \max\{\Psi_1, \Psi_2, \Psi_3, \Psi_4\}$ with (Ψ_2, Ψ_3, Ψ_4) defined in appendix. The definition implies $\Psi_2 \geq \Psi_1$. I also have that the condition on γ is stricter than $\gamma\kappa\rho \leq 1$ in proposition 1 because $\kappa\rho(1-\kappa\rho)\rho\frac{\sigma-1}{\sigma} < 1$. In the no-screening case when either ψ or κ goes to ∞, the results for $(l, q, I_X, r, w, \Lambda, [\Delta/\check{l}], [V/\check{l}], [u'/\Delta])$ are valid for the global saddle path equilibrium and for any (γ, κ, ρ).

I distinguish between growing firms converging from below ($h < h_S$) and shrinking firms converging from above ($h > h_S$). A gradual increase in the number of workers, the screening cutoff[25] and worker average ability is optimal because of adjustment costs convexities and the fact that screening compounds per-worker recruiting costs. Nevertheless, firms grow faster when h is smaller and they are further away their optimal stationary state because of diminishing returns that make the opportunity costs of delaying recruitment higher. This property is illustrated by the fact that both the hiring rate $[{}^{\Delta}/_{\check{l}}]$ and the vacancy rate $[{}^{V}/_{\check{l}}]$ decrease with h. Obviously, growing firms do not fire without screening ($s = s_S$).

Shrinking firms lower their screening cutoff down to the stationary level and do not adjust it thereafter.[26] I only consider cases where the cutoff and employment are together above or below their stationary level ($(\alpha_c - \alpha_S)(l - l_S) \geq 0$), implying that shrinking firms aim at decreasing both employment and the screening level. The restrictive assumptions about chosen cutoffs ($\alpha_\infty \geq \alpha'_c \geq min(\alpha_c, \alpha'_{c,\infty})$) then implies that shrinking firms choose $\alpha'_c = \alpha_S$ once and for all.

Firm employment can shrink either by attrition when firms let the hiring rate fall below the quit rate ($[{}^{\Delta}/_{\check{l}}] < s_S$), or by firing without screening $s > s_S$. The higher h and the further away a firm is from its stationary state, the larger the opportunities for cost savings: reducing employment has a small impact on production because of diminishing returns whereas it has a substantial cost-cutting effect because of the reduction in payments of the wage outside option component ($\hat{w}_o l$). Therefore, firing increases and hiring falls with firm size.

Two opposing effects compete to determine the evolution of wage offers and rejection rates. Small firms have lower screening and rejection rates and do not need to compensate workers as much. Yet, small firms also want to offer higher wages to attract more workers per vacancy and to speed growth. In the process, more workers will not get an interview because of matching frictions. When ψ is high enough, the latter effect dominates. Wage offers, the searcher-vacancy ratio and the rejection rate vary similarly: all decrease with firm size.

[25] This result implies that the restrictive assumption on α'_c in the firm problem (15) is not binding when ϕ and Υ are constant (i.e. in steady states) as firms start small.

[26] In the general equilibrium considered in the next section, this only happens right after the unexpected trade reform, which is referred to as "the first period".

2 General equilibrium

2.1 Endogenous entry and exit and markets clearing

Periods are once again indexed by t, t_0 is the date of the unexpected trade reform implementation, and firm age is indexed by a.

2.1.1 Endogenous entry and exit

Only a fraction of the new firms ($a = 0$) created every period, if any, decides to become active. When their productivity is revealed upon entry, the new firms that are not productive enough to generate positive profits decide to exit. The productivity distribution of the remaining entrants is Pareto with cutoff $\underline{x}_{0,t}$: $\delta_{0,t}(x) = 1$ if $x < \underline{x}_{0,t}$, and $\delta_{0,t}(x) = \delta_S$ otherwise. The marginal entrants' value, with productivity $\underline{x}_{0,t}$, is null:[27]

$$G(\alpha_{c,0,t}, l_{0,t}, 0, \underline{x}_{0,t}; \mathscr{A}_t) = 0 \tag{ZCP}$$

For all periods except when the trade reform is unexpectedly implemented, no firm with age $a > 0$ exit (see section 1.2.1), implying: $\delta_{a,t}(x) = \delta_S$ for all $x \geq \underline{x}_{a,t}$ with $\underline{x}_{a,t} = \underline{x}_{a-1,t-1}$ when $t \neq t_0$.

However at $t = t_0$, the trade reform is unexpected and may result in voluntary exits. Some firm values can become negative if the present value of the sum of future revenues become greater than the sum of minimum unavoidable costs, including wage commitments. I assume that these firms either exit or renegotiate wage commitments with their incumbent workers.[28] All workers at a renegotiating firm are assumed to get the same percentage cut.

For firms with negative value at $t = t_0$, I distinguish two cases. When a firm would still have a negative value after cutting all wage premia to zero, I assume that it exit. Otherwise, there are wage cuts that can both restore the firm value to non-negative levels while preserving positive premia over the workers outside option. In this case, cuts are a Pareto improvement over firm exit. I assume that all the bargaining power is on the side of workers: therefore, wage cuts $cut_{a,t_0}(x)$, if any, are just enough to bring the firm value from negative to null. This is a conservative

[27] This condition is the analogue of the zero cutoff profit condition in M03.

[28] Both options are a violation of the commitments made to workers at the time of their hiring. When a firm exits with probability $\delta = 1 > \delta_S$, all employees are laid-off and lose promised expected wage premia. I assume these deviations from the norm are justified by the exceptional event.

hypothesis in the sense that it minimizes wage cuts at already low-paying firms and therefore minimizes the change in inequality that results from renegotiation.[29] Formally, I assume:

$$
\begin{array}{ll}
\delta_{a,t_0}(x) = \delta_S \text{ , } cut_{a,t_0}(x) = 0 & \text{if } G(\alpha_{c,a,t_0}, l_{a,t_0}, B_{a,t_0}, x; \mathscr{A}_{t_0}) \geq 0 \\
\begin{cases} \delta_{a,t_0}(x) = \delta_S \text{ , } B_{a,t_0} \geq cut_{a,t_0}(x) > 0 \\ G(\alpha_{c,a,t_0}, l_{a,t_0}, B_{a,t_0} - cut_{a,t_0}(x), x; \mathscr{A}_{t_0}) = 0 \end{cases} & \text{if } \begin{cases} G(\alpha_{c,a,t_0}, l_{a,t_0}, B_{a,t_0}, x; \mathscr{A}_{t_0}) < 0 \\ G(\alpha_{c,a,t_0}, l_{a,t_0}, 0, x; \mathscr{A}_{t_0}) \geq 0 \end{cases} \\
\delta_{a,t_0}(x) = 1 & \text{if } G(\alpha_{c,a,t_0}, l_{a,t_0}, 0, x; \mathscr{A}_{t_0}) < 0
\end{array}
\tag{22}
$$

For any age $a > 0$, the above conditions are a full partition of the productivity space because firm values are increasing in productivity. Therefore at t_0, we can define the cutoff $\underline{x}_{a,t_0} = \min\{x | G(\alpha_{c,a,t_0}, l_{a,t_0}, 0, x; \mathscr{A}_{t_0}) \geq 0\}$ between active and inactive firms of age a.

I assume free entry of new firms. In any period t, it implies that the expected value of entering firms is at most equal to the sunk cost of entry in equilibrium:

$$
(1-\delta_S) \int_{\underline{x}_{0,t}}^{\infty} G(\alpha_{c,0,t}, l_{0,t}, 0, x; \mathscr{A}_t) dF(x) \leq f_e \text{ and } M_{0,t} \geq 0 \text{, with at least one equality} \tag{FE}
$$

Equality for the first condition is obtained when the mass $\frac{M_{0,t}}{1-\delta_S}$ of firms that enters[30] is strictly positive. Subsequently, the mass of incumbent firms evolves subsequently according to

$$
M_{a,t}(x) = (1 - \delta_{a,t}(x)) M_{a-1,t-1}(x) \tag{23}
$$

which simplifies to $M_{a,t} = (1-\delta_S) M_{a-1,t-1}$ for $x \geq \underline{x}_{a,t}$, and $M_{a,t} = 0$ otherwise.

2.1.2 Aggregate consumption over time and markets clearing

Aggregate consumption and the interest rate are determined using a representative agent because I assume nested CES preferences and complete financial markets.

I assume that each worker has an equal participation in a national fund as in Chaney (2008): [7] profits are equally rebated among workers of the two symmetric

[29] In general, the importance of nominal wage cuts in the data is a source of debate. The wage cuts predicted by the model are in real terms. Furthermore, there is in the literature some supportive evidence of wage cuts, even nominal ones, under special circumstances (surveyed in Akerloff et al. (1996) [1]) such as when firms experience particularly harsh difficulties as would be the case here.

[30] The $\frac{M_{0,t}}{1-\delta_S}$ entrants are hit by the exogenous exit shock: at most a remaining mass $M_{0,t}$ is active.

economies. I assume that the representative consumer only consumes the differentiated good. Intra-temporal preferences over varieties are as posited in subsection 1.2.1. Inter-temporal preferences are captured by a utility function with constant elasticity of inter-temporal substitution IES and a discount parameter β. In any period t and taking as given the sequences of total future revenues and aggregate price indexes $(\mathscr{Y}_{t'}, \mathscr{P}_{t'})_{t'=t..\infty}$, the representative agent utility function is:

$$\mathscr{U}_t = \mathbb{E}_t \sum_{t'=t}^{\infty} \beta^t \frac{1-\mathscr{C}_t^{1-IES}}{IES-1} \qquad \text{subject to } \mathbb{E}_t \sum_{t'=t}^{\infty} \frac{1}{R_{t,t'}} \left(\mathscr{P}_{t'}\mathscr{C}_{t'} - \mathscr{Y}_{t'}\right) = 0 \tag{24}$$

The equilibrium interest rate is thus given by $R_{t,t+1} = \frac{\mathbb{E}_t}{\beta}\left(\frac{\phi_{t+1}}{\phi_t}\right)^{\sigma IES}\left(\frac{\mathscr{P}_{t+1}}{\mathscr{P}_t}\right)^{1-\sigma IES}$, a standard Euler equation where I use the competition index $\phi_t \equiv \mathscr{C}_t^{\frac{1}{\sigma}} \mathscr{P}_t$.

Total population POP must be equal to the sum of workers "attached" to every firm of every generation of age a. Each worker is directly or indirectly "attached" to a firm, either because he or she works there ($l'_{a,t}$), or had unsuccessfully applied there ($u'_{a,t}$), or produces the services purchased by the firm (operations $f_d + I_{X,a,t} f_X$, technology support $\frac{c_s}{\psi}\overline{\alpha}_{a,t}^{\psi}$, vacancy posting $\xi A_{a,t}$). The dependence of firm variables on $(x;\mathscr{A}_t)$ is implicit. The labor market clearing condition is:

$$POP = M_{0,t} f_e + \sum_{a\geq 0} \int_{\underline{x}_{a,t}}^{\infty} M_{a,t} \left[u'_{a,t} + l'_{a,t} + f_d + I_{X,a,t} f_X + \xi A_{a,t} + \frac{c_s}{\psi}\overline{\alpha}_{a,t}^{\psi}\right] dF(x) \tag{LMC}$$

Total consumption equals total revenues and these amount to firm revenues because firm costs are eventually paid to workers in services and because profits are rebated to workers. Hence, the good market clearing and the revenues equations are as follows:

$$\mathscr{P}_t^{1-\sigma}\phi_t^{\sigma} = \mathscr{C}_t = \mathscr{Y}_t = \sum_{a\geq 0} \int_{\underline{x}_{a,t}}^{\infty} M_{a,t} r_{a,t}\left(x;\phi_t,\Upsilon_t\right) dF(x) \tag{GMC}$$

Adjustments in the real wages occur through adjustments of the aggregate price[31] $\mathscr{P}$ because nominal wages are pinned down by productivity in the service sector.

[31] The aggregate price is $\mathscr{P} \equiv \left[\sum_{a\geq 0} \int_{\underline{x}_{a,t}}^{\infty} M_{a,t}(x)\left((p_{a,t}^D(x))^{1-\sigma} + I_{X,a,t}(x)(p_{a,t}^F(x))^{1-\sigma}\right) dF(x)\right]^{\frac{1}{1-\sigma}}$ and consistent with the standard definition.

2.1.3 General equilibrium definition

I consider the case of an economy that is possibly hit by an unexpected one-off reduction in trade costs at the beginning of t_0 (the trade reform).

Given initial conditions in $t_0 - 1$ and an unexpected change in trade costs (if any) from $(\tau_0, f_{X,0})$ to $(\tau_\infty, f_{X,\infty})$ at t_0, a general equilibrium is a sequence of firm policy functions $\{\delta_{a,t}, \alpha'_{c,a,t}, l_{a,t}, \Delta_{a,t}, V_{a,t}, w_{a,t}, cut_{a,t}, I_{X,a,t}\}$ for every $t \geq t_0$, $a \geq 0$ and $x \sim F()$ and a sequence of aggregate variables $\{\phi_t, \mathscr{P}_t, M_{0,t}\}_{\{t \geq t_0\}}$ such that:

1. at $t = t_0$, firms possibly negotiate wage cuts or exit according to (22);

and after renegotiations, if any, and at all times $t \geq t_0$:

2. the unemployed make optimal searching decisions taking vacancy characteristics as given and ruling out any future wage renegotiation (10)-(11)
3. firms choose policy functions to maximize their value as described in (15) taking aggregate variables as given; they comply with wage commitments;
4. the free entry (FE) and the zero-cutoff (ZCP) conditions are satisfied;
5. the interest rate $R_{t,t+1}$ follows the Euler equation related (24);
6. the product and labor markets clear (conditions (GMC)-(LMC) are satisfied).

2.2 Steady state equilibrium and the long run effects of trade

In this subsection, I compare steady states with different levels of trade costs. I define steady states as the general equilibria with constant trade costs where all aggregate variables $(\phi, \mathscr{P}, M_0)$ are constant. I drop the subscript t. I restrict my attention to equilibria in which the mass of entrants is strictly positive.

To derive some analytical results, I consider two special cases with restrictive assumptions. Interestingly, results obtained in these special cases hold in the more general settings considered in the numerical explorations of next section.

2.2.1 Steady state special cases

The two special cases considered have four assumptions in common. First, I assume that there is no impatience as in M03, meaning that $\beta = 1$. Second, I assume that productivity draws are not too dispersed, and precisely that $\theta \geq \max\left\{\frac{\sigma-1}{\sigma}\varepsilon_\pi, \left(\varepsilon_\pi \frac{\sigma-1}{\sigma}\left(2^{\frac{\varepsilon_\pi}{\sigma}} - 1\right) - \varepsilon_l \frac{\sigma-1}{\sigma}\left(2^{\frac{\varepsilon_\pi}{\sigma}} - 2^{\frac{\varepsilon_\pi - \varepsilon_l}{\sigma}}\right)\right) / \left(2^{\frac{\varepsilon_\pi - \varepsilon_l}{\sigma}} - 1\right)\right\}$. Third, I assume that the starting number and characteristics of workers at new firms is proportionate to the firm characteristics in the stationary equilibrium, meaning that $l_0(x;\phi,\Upsilon) = \iota . l_S(x;\phi,\Upsilon)$ and $\alpha_{c,0}(x;\phi,\Upsilon) = \iota . \alpha_S(x;\phi,\Upsilon)$ for some parameter $\iota \in (0,1)$. Fourth, I assume firms can only export from the beginning or not at all.

Each special case has a specific additional fifth assumption. In the first case, henceforth **SC1**, I consider the case of no-screening where ψ or κ goes to infinity, while in the second case, henceforth **SC2**, and as in HIR10, I consider the case where the outside option of workers is null, meaning that $w_o = 0$.[32]

Now that I assume that firms decide once and for all at birth whether to export, their decision only depends on their productivity and steady state aggregate variables. Exporting firms (with $I_X = 1$) are those with a productivity draw x greater or equal than the export cutoff $\underline{x}_X(\phi,\Upsilon)$. For every firm, I then define "effective productivity" as $X = x^{\frac{\sigma-1}{\sigma}} \left[1+(\Upsilon-1) I_X(x;\phi,\Upsilon)\right]^{\frac{1}{\sigma}} \phi$.

Lemma 1 (Log-linearity of firm variable in the two special cases)
Firm variables are log-linear in X: for any $z \in \{\alpha_c', l, \Delta, w, \pi, r, u, A, B\}$*, there are coefficients* (z_a^*, ε_z) *independent of* (x,ϕ,Υ,R) *such that* $z_a(x;\phi,\Upsilon,R) = z_a^* . X^{\varepsilon_z}$*. The hiring rates,* $[\Delta/\check{l}]_a$*, and firing rates, s, are independent of* (x,ϕ,Υ,R,X) *and* $s = s_S$*.*
*-**SC1**:* $\varepsilon_\alpha = \varepsilon_w = 0$*, and* $\varepsilon_\Delta = \varepsilon_l = \varepsilon_r = \varepsilon_\pi = \varepsilon_A = \varepsilon_B = \varepsilon_u = \left(1-\rho\frac{\sigma-1}{\sigma}\right)^{-1}$
*-**SC2**:* $\varepsilon_\alpha = \frac{1}{\psi\left(1-\rho\frac{\sigma-1}{\sigma}\right)-\frac{(\sigma-1)(1-\kappa\gamma\rho)}{\sigma}}$*,* $\varepsilon_w = \kappa\gamma\varepsilon_\alpha$*,* $\varepsilon_\Delta = \varepsilon_l = (\psi-\kappa\gamma)\varepsilon_\alpha$*,* $\varepsilon_r = \varepsilon_\pi = \varepsilon_u = \varepsilon_A = \varepsilon_B = \psi\varepsilon_\alpha$

Lemma 1 provides an extension of proposition 1 corollary 1: the elasticities of firm variables with respect to x, ϕ and Υ at any given age a are the same as in the firm stationary states. Holding age constant, more productive firms have more workers, with better ability and higher wages in the screening case,[33] sell more and make more profits.

The hiring rates are independent of productivity for two reasons. First, there are no economies of scale in hiring or firing, and hiring costs depends on the relative increase in employment. Second, hiring needs at any given age are the same for firms with different productivity because more productive firms are assumed to start with

[32] Equation (5) implies that it can be so when the dis-utility of working in services is high enough.
[33] Wage offers are independent of productivity when there is no screening as in the **SC1**. In that case, wages only vary with hiring rates, and therefore only with age.

more and better workers and are thus at the same distance from their stationary state than smaller less productive firms. Note that Lemma 1 implies that firm variable growth rates are independent from productivity and trade costs.

In steady state, firms starting small never grow to exceed their stationary states. Therefore, in the absence of adverse shocks, they never need to fire.

2.2.2 Steady state welfare decreases with trade costs in the special cases

In steady states, R equals $1/\beta$ which is one by assumption in special cases. Also, the mass of firms with age a simplifies to $M_a = (1-\delta_S)M_{a-1} = (1-\delta_S)^a\,(1-F(\underline{x}))\,M_0$ because firms voluntarily exit only at birth, and the cutoff is the same for all ages, $\underline{x}_a = \underline{x}_0$. Thus, I drop $\underline{x}$ subscripts. Similarly, the export cutoff is defined as $\underline{x}_X$.

For all firm variables z, I define $\bar{z}(x;\phi,\Upsilon)$ as the discounted average of z over any firm x lifespan. I prove in appendix the following.

Lemma 2 (Log-linearity and equivalence properties of firm lifespan averages) *In both special cases and for any $z \in \{\alpha_c', l, \Delta, w, \pi, r, u, A, B\}$, there are coefficients (z_*^*, ε_z) independent of (x,ϕ,Υ,R) such that $\bar{z}(x;\phi,\Upsilon,R) = z_*^*.X^{\varepsilon_z}$, where ε_z are the same as in Lemma 1. Moreover at any time t, the averages across age of $z_a(x)$ among existing firms with productivity x are equal to firm x lifespan averages, $\bar{z}(x)$.*

$$\bar{z}(x;\phi,\Upsilon,R) \equiv \frac{\sum_{a\geq 0}\left((1-\delta_S)/R\right)^a z_a(x)}{\sum_{a\geq 0}\left((1-\delta_S)/R\right)^a} = \frac{\sum_{a\geq 0} M_a z_a(x)}{\sum_{a\geq 0} M_a} = z_*^*.X^{\varepsilon_z} \tag{25}$$

The first part of the lemma makes use of the log-linearity property to take averages across one dimension of firm heterogeneity, namely age, independently of the other dimension, namely productivity. The second part stems from the assumption of no impatience ($R = 1$).

The special cases log-linearity property allows me to simplify the conditions that characterize the equilibrium by eliminating references to age. I obtain:

$$\textbf{ZCP}^{34} \qquad \pi_*^* \underline{x}^{\frac{\sigma-1}{\sigma}\varepsilon_\pi} \phi^{\varepsilon_\pi} - f_d = 0 \quad \text{and} \quad \pi_*^* . \left(\underline{x}_X^{\frac{\sigma-1}{\sigma}} \phi \right)^{\varepsilon_\pi} \left(\Upsilon^{\frac{\varepsilon_\pi}{\sigma}} - 1 \right) - f_X = 0$$

$$\textbf{FE} \qquad \frac{1-\delta_S}{\delta_S} \int_{\underline{x}}^{\infty} \pi_*^* . X^{\varepsilon_\pi} - f_d - f_X I_X dF(x) = f_e$$

$$\textbf{LMC} \quad POP = \frac{M_0 f_e}{1-\delta_S} + \frac{M_0}{\delta_S} \int_{\underline{x}}^{\infty} (u_*^* + \xi A_*^*) X^{\varepsilon_\pi} + \frac{c_s}{\psi} \alpha_*^{*\psi} X^{\psi \varepsilon_\alpha} + (l_*^* - \Delta_*^*) X^{\varepsilon_l} + f_d + I_X f_X dF(x)$$

$$\textbf{GMC} \qquad \mathscr{P}^{1-\sigma} \phi^\sigma = \mathscr{Y} = \frac{M_0}{\delta_S} \int_{\underline{x}}^{\infty} r_*^* . X^{\varepsilon_\pi} dF(x) \qquad \left(\text{with } X \equiv x^{\frac{\sigma-1}{\sigma}} \phi \left[1 + (\Upsilon - 1) I_X \right]^{\frac{1}{\sigma}} \right)$$

In solving for the equilibrium, the cutoffs and the competition index ϕ first are obtained by combining the **ZCP** and **FE** conditions. Then, the mass of entrants clears the labor market (**LMC**) and the price index the good market (**GMC**).

Welfare in the steady state is defined as consumption per capita: $\mathscr{W} \equiv \mathscr{C} / POP$.

Proposition 2 (Comparisons across special cases steady states)
Welfare decreases with the level of the fixed or variable trade costs, and so does the mass of entrants, average firm productivity and the average level of screening. Conversely, the price index increases. When trade costs increase, "effective productivity", X, falls for domestic firms but increases for exporters.

In steady state special cases, gradual firm growth plays no significant role and the logic behind the standard M03 and HIR10 models applies. As trade barriers fall, exporters make more profits from foreign sales and grow by hiring and screening better workers, implying that expected profits for entrants improve. Entrepreneurs decide to create more firms, increasing the intensity of competition as reflected by a decrease in the price index. In equilibrium, free entry implies that the expected gains from better export opportunities are offset by a lower entry probability driven by the intensification of competition: the entry cutoff rises. Moreover, a lower price index induces a reallocation of labor away from low productivity low screening domestic firms. Thus, average firm and worker productivity increase. New varieties from abroad compensate the reduction in domestic varieties, and welfare rises.

[33] I assume that the marginal entrant does not export. Computation details are in appendix.

2.2.3 Connections with respect to the workhorse Melitz model

Corollary 1 to proposition 2 (Equivalence results with Melitz (2003)[25])
*Consider two models, the **SC1** dynamic model with $\rho = 1$ and with any firm-productivity distribution, and a counterfactual M03 static trade model where entering firms draw productivity from the same distribution. All common parameters are the same in the two models except for an industry-wide productivity shifter Θ_1, a service sector productivity shifter Θ_2, and a population shifter Θ_3.*[35]

Then, aggregate outcomes $(\mathscr{C}, \mathscr{P}, M_0)$ and the elasticity of welfare with respect to openness measures (Υ or $\tilde{f}_d/f_X$) are exactly identical in both models. Moreover, if firm productivity draws are Pareto distributed, the formula developed in Arkolakis et al. (2012)[3] applies: welfare gains from trade are proportionate to the share of domestic expenditure on domestic goods ($\mathscr{E}$) in log terms ($\widehat{\mathscr{W}} = \widehat{\mathscr{E}}^{-1/\theta}$).

The corollary implies that central steady state results of the literature on total gains from trade can be extended to dynamic settings. Log-linearity breaks all links between productivity and growth rates and allows for a mapping between firm variables in the static model and the firm lifespan averages in the dynamic model. The evolution of firms following their creation can then be sufficiently represented by static average values as in M03, and the equivalence results follow.

The results could be generalized to other dynamic models with log-linearity. Similar equivalence results would hold if one alternatively assumes an ad-hoc evolution of firm variables, namely any $\{z_a^*\}_{a \geq 0}$ sequence, as long as this evolution (the sequence) is common to all firms and remains independent from aggregate variables and firm productivity. In the present model, the evolution is micro-funded by labor search and screening costs.

2.2.4 Distributional effects across sectors in steady state special cases

In this section, I start studying the steady state distribution of revenues across sectors by considering the sectoral allocation of labor. I consider the differentiate sector average wage, $\overline{W}$, and the sector unemployment rate at the end of each period,

[35] I define $\Theta_1 = \frac{\sigma}{\sigma-1}\frac{\kappa-1}{\kappa\alpha_{min}} {r_*^*}^{1/(\sigma-1)}$, $\Theta_2 = \frac{1}{\sigma}\frac{r_*^*}{r_*^* - \xi A_*^* - w_o l_*^* - B_*^*}$, and $\Theta_3 = \frac{r_*^*}{r_*^* - w_o l_*^* - B_*^* + l_*^* + u_*^* - \Delta_*^*}$ such that firm productivities are given by $x\frac{\kappa\alpha_{min}}{\kappa-1}\Theta_1$, the fixed entry cost is $\Theta_2 f_e$, the operation cost is $\Theta_2(f_d + \frac{c_s}{\psi}\alpha_{min}^{\psi})$, the export cost is $\Theta_2 f_X$, and population is $\Theta_3 POP$. See appendix for details.

U. I define U as the ratio of unmatched job-searchers over the sum of job-searchers and employed workers.[36] I also define LF_p as the share of workers tied to the differentiated sector (employed or seeking a job).

$$\overline{W} \equiv \sum_a \frac{M_a}{(1-F(\underline{x}))} \int_{\underline{x}}^{\infty} B'_a(x) dF(x) \Big/ \sum_a \frac{M_a}{(1-F(\underline{x}))} \int_{\underline{x}}^{\infty} l'_a(x) dF(x) \tag{26}$$

$$U \equiv \frac{\sum_a M_a/(1-F(\underline{x})) \int_{\underline{x}}^{\infty} (u_a(x) - \Delta_a(x)) \, dF(x)}{\sum_a M_a/(1-F(\underline{x})) \int_{\underline{x}}^{\infty} (u_a(x) + (l_a(x) - \Delta_a(x))) \, dF(x)} \tag{27}$$

$$LF_p \equiv \sum_a \frac{M_a}{(1-F(\underline{x}))} \int_{\underline{x}}^{\infty} \left(u_a(x) + l'_a(x) - \Delta_a(x)\right) dF(x) \Big/ POP \tag{28}$$

I prove the following proposition in appendix.

Proposition 3 (Steady state comparison of labor allocation across sectors)
*In **SC2**, $\overline{W}$, U and $1-LF_p$ decrease with the level of trade costs, fixed or variable. These are constant in **SC1**.*

The increase in unemployment and the average wage as an economy opens result from two effects. First, increased competition induces a reallocation towards more productive firms with a higher level of screening and with vacancies with longer queues. Second, new sales opportunities abroad provide incentives for exporters to screen even more and post vacancies with even longer queues. More screening and longer queues on average are compensated by a higher average wage. By contrast, unemployment and the average wage remain constant when there is no screening.

The same two effects are at play in determining the relative size of sectors because the share of service workers is related to screening intensity. Firms that screen more have to equip their production workers with a better support technology that requires more service workers to produce. Firms that screen more also produce more with a smaller share of more able production workers. Hence, both the trade-induced reallocation towards more productive screening-intensive firms and increased screening at exporters contribute to increase the indirect employment of service workers relative to the direct employment of production workers.

[36] Similar results are obtained using the ratio of unmatched job-searchers over total population as an alternative definition of unemployment. See appendix for details.

2.2.5 Steady state wage inequality within the differentiated sector

Different cohorts of hires have different wages within and across firms. I define as a cohort all the workers that were hired in the same period and are still employed at the same firm. By construction, workers in the same cohort have the same wage and tenure (measured in number of periods). I index cohorts by their tenure c.

I show by iteration (with details in appendix) that the size $l'_{a,c}(x)$, cutoff $\alpha'_{c,a,c}(x)$ and wage $w_{a,c}(x)$ of workers in every cohort are log-linear in X because screening $\alpha'_{c,a}(x)$ and recruitment $\Delta_a(x)$ are so. In turn, I show that the average wage at a firm x of age a, denoted by $\hat{w}_a(x)$, is also log-linear in X. I then show that log-linearity implies that the wage dispersion within any firm as measured by the coefficient of variation, $cv_a(x;\phi,\Upsilon,R)$, is independent from (x,ϕ,Υ,R) and only depends on age.

$$\hat{w}_a(x;\phi,\Upsilon,R) \equiv \frac{\sum_{c=0}^{a} l'_{a,c}(x) w_{a,c}(x)}{\sum_{c=0}^{a} l'_{a,c}(x)} = \hat{w}_a^*.X^{\varepsilon_w} \tag{29}$$

$$cv_a(x;\phi,\Upsilon,R) \equiv \frac{1}{\hat{w}_a(x)}\sqrt{\frac{\sum_{c=0}^{a} l'_{a,c}(x)\left(w_{a,c}(x)-\hat{w}_a(x)\right)^2}{\sum_{c=0}^{a} l'_{a,c}(x)}} = cv_a^* \tag{30}$$

Wage dispersion within firms only stems from differences in historical growth rates because the other wage determinants, productivity and screening, are the same for all workers in a firm at a given period. Moreover, lemma 1 implies that firm growth rates are independent of productivity and aggregates. Therefore, productivity, screening and aggregates simply move all wages within a firm in the same way and do not affect within-wage dispersion.

I define aggregate within-firm inequality $\overline{CV}_W$ as the weighted average of within-firm coefficient of variations and obtain that it is independent from trade costs.

$$\overline{CV}_W \equiv \frac{\sum_a \frac{M_a}{(1-F(\underline{x}))}\int_{\underline{x}}^{\infty} l'_a(x) cv_a(x) dF(x)}{\sum_a \frac{M_a}{(1-F(\underline{x}))}\int_{\underline{x}}^{\infty} l'_a(x) dF(x)} = \frac{\sum_a (1-\delta_S)^a l_a^* cv_a^*}{\sum_a (1-\delta_S)^a l_a^*} \tag{31}$$

Measuring total inequality by the coefficient of variation of all wages in the sectors,

$$\overline{CV}_T \equiv \frac{\sum_a \frac{M_a}{(1-F(\underline{x}))}\int_{\underline{x}}^{\infty}\sum_{c=0}^{a} l'_{a,c}(x)\left(w_{a,c}(x)-\overline{W}\right)^2 dF(x)}{\sum_a \frac{M_a}{(1-F(\underline{x}))}\int_{\underline{x}}^{\infty}\sum_{c=0}^{a} l'_{a,c}(x) dF(x)} \tag{32}$$

and between-firm inequality as the difference between the measures of total- and within-wage inequality, $\overline{CV}_B \equiv \overline{CV}_T - \overline{CV}_W$, I prove the following in appendix.[37]

Proposition 4 (Steady state comparison of sectoral wage inequality measures) *(i) In both special cases, $\overline{CV}_W$ is independent of trade costs, and (ii) both $\overline{CV}_T$ and $\overline{CV}_B$ are the same when all firms export as when no firms export.*
*(iii) In **SC1**, $\overline{CV}_T$ and $\overline{CV}_B$ are independent from trade costs, while (iv) in **SC2**, both $\overline{CV}_T$ and $\overline{CV}_B$ when some but not all firms export is strictly greater than in autarky. Moreover in **SC2**, an increase in the fraction of exporting firms raises sectoral wage inequality when the fraction of exporting firms is sufficiently small, and reduces wage inequality when the fraction of exporting firms is sufficiently large.*

A reduction in trade costs amounts to a productivity boost for exporters and a productivity reduction for domestic firms (proposition 2). In **SC1**, the absence of screening severs the firm productivity-wage link: the level of trade does not affect wage dispersion. In **SC2**, trade generates a wage differential between domestic and exporting firms. When trade costs are high, only a small fraction of workers are employed at exporters and benefit from this wage differential: a decrease in trade costs raises the share of workers at exporters and thereby increases wage dispersion. When trade costs are low, the majority of workers are employed at exporters and a trade cost reduction implies that even more workers benefit from the wage differential: wage dispersion decreases. When all firms export, all workers benefit from the same exporter differential and wage dispersion is similar to the case when no firms export. Hence, changes in inequality operates in the same as in HIR10.

Furthermore, I establish in appendix an equivalence between variations with trade costs of wage and employment in HIR10 and firm lifespan average wage and employment in **SC2**. As for the other equivalence results, this one can generalized to any dynamic model with log-linearity in X as defined in lemma 1.

In conclusion, the two equivalence results have two important consequences: first, they highlight the fact that dynamic features do not necessarily matter for steady state comparisons, and second, they mean that all deviations from steady states observed during a transition are also deviations from standard results.

[37] I show in appendix that the results on between-firm inequality are robust to the use of alternative and perhaps more standard definitions.

2.3 The equilibrium path after a reduction in trade costs

In this subsection, I study the transition path from a steady state characterized by $(\phi_0, \mathscr{P}_0, \tau_0, f_{X,0})$ after a once and for all reduction in trade costs to $\tau_\infty \leq \tau_0$, $f_{X,\infty} \leq f_{X,0}$, with at least one strict inequality. Aggregates in the new future steady state are indexed by ∞. To derive some analytical results, I assume that $IES\sigma = 1$, that marginal entrants never export, and that the reduction in trade costs is "not too big".[38] Individual firm equilibrium paths obtained analytically under these assumptions are confirmed by the next section numerical results without restrictions.

I show in appendix that there is a transition path equilibrium where the market competition index immediately adjusts to the new steady state value, meaning that $\phi_t = \phi_\infty$ and $R_{t,t'} = 1/\beta$ for all $t \geq t_0$. From the firms' perspective, the transition is a quasi steady state. I also show that intensified competition is such that $\phi_\infty \leq \phi_0$, and that new market opportunity are reflected by the increase Υ such that $\phi_\infty \Upsilon_\infty^{\frac{1}{\sigma}} \geq \phi_0 \Upsilon_0^{\frac{1}{\sigma}}$.

2.3.1 The evolution of individual firms and their wage policies

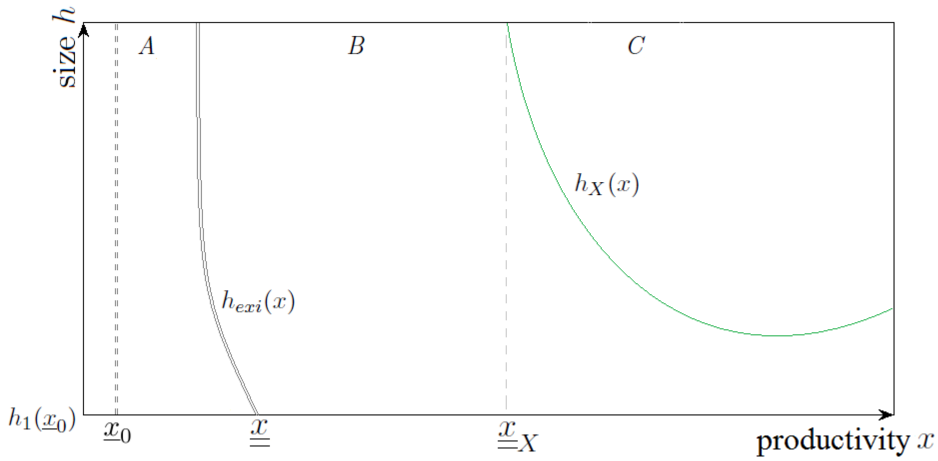
Because the economy is in a quasi steady state, new entrants behave as any firm in the future steady state. They start small with low screening levels and gradually expand according to the policy functions of proposition 1 and its corollaries.

The firms that entered before the reform make decisions that can be radically different from what they would have done without the reform. Growth patterns are characterized in the following proposition with cutoff curves illustrated in figure 2.

Proposition 5 (The saddle path of firms born before the reform)
At t_0, there is a partition of existing firms in three groups, A, B and C, such that $A = \{x \leq \underline{\underline{x}}, h < h_{exi}(x)\}$ with $\partial h_{exi}/\partial x \leq 0$, $C = \{x \geq \underline{\underline{x}}_X, h \geq h_X(x)\}$, and $B = \{(x \geq \underline{x}_0, h \geq h_1(\underline{x}_0; \phi_0, \Upsilon_0)) \notin A \cup C\}$ and such that (i) group A firms exit at t_0, (ii) group C firms will export in their new future stationary state, while surviving group B firms will not. Furthermore, (iii) group B firms lower their screening cutoff and shrink, or increase employment and their screening cutoff less rapidly than they would have without the reform, and (iv) group C firms increase employment and

[38] The later assumption is made to ensure that the mass of entrants is always strictly positive, thereby implying that the free entry condition is always binding, as well as to ensure that the local results of proposition 1 apply.

Figure 2: Illustrative diagram of the firms partition in proposition 5.



Group A firms that exit are to the left of the plain double gray line. Group C firms that will export in their new stationary state are to the right and above the plain single green line while group B firms that won't are below and to the right of it.

their screening cutoff, and among them, those who immediately export do it faster than they would have without the reform.

The group A firms that exit after at the reform date are the smaller and less productive firms with negative values and for which the largest negotiable wage cuts cannot bring their value back to non-negative territory. These firms don't export and don't benefit from the trade cost reduction. On the contrary, increased competition from openness decrease their present and future revenues. They do not find profitable anymore to pay for the fixed cost of operation or to pay for the adjustment costs that would allow them to grow to a size where it would be profitable to pay for the fixed operation cost.

The group C firms are the more productive and larger firms that find it profitable to pay the hiring and screening adjustment costs to grow and benefit sustainably from the new sales opportunity abroad. Only the firms that are productive enough to export in their stationary state belong to this group, but being productive is not a sufficient condition. Adjustment costs can be prohibitive for the firms that are not large enough in t_0 to profitably export in that period. Small firms that could grow to a size where export would be profitable may instead save on adjustment costs, remain small and never export. Conversely, the larger firms in group C face low enough adjustment costs and decide to grow to export immediately or in the future.

Group C firms are below their new stationary state and start to increase employment and screening immediately after the reform. Indeed, their new stationary size and screening level with the reform are larger than those without the reform because of improved market opportunities abroad. Moreover, the firms that are productive and large enough to export immediately after the reform are unambiguously better off because they immediately benefit from the lower trade costs: they immediately hire more and screen more than what they would do without the reform.

The group B firms are the firms that are productive and large enough to survive but not so much as to become stationary exporters. Their new stationary size and screening level with the reform are lower than those without the reform. I distinguish two subgroups. First, the old firms that grew close enough or exactly to their pre-reform stationary state are now above their new stationary state. By assumption, these firms immediately lower the screening level to the new stationary level. They also reduce employment either by recruiting less workers than the number of exogenous quits (attrition) or by actively firing some workers. Second, the young firms that had little time to grow are still below their stationary state. These firms keep increasing employment and the screening level, but less than what they would do without the reform.

2.3.2 Wage comparisons between the transition and steady states

Proposition 5 has the following implications (proved in appendix):

Corollary 1 to proposition 5 (Wage policy of firms born before the reform)
At t_0, (i) the firms that cut existing wages, if any, are less productive and smaller domestic firms. In comparison with what they would do without the reform, (ii) group B firms offer lower wages to new hires and, if still growing, raise the wage of their incumbent workers less and slowlier (iii) the group C firms that immediately export offer higher wages to new hires and raise more and faster the wage of the ones they retain.

Among workers hired before the reform, the ones (if any) that get wage cuts at t_0 are employed at the surviving firms whose value drop the most. Assumption 22 implies that these firms are smaller and less productive surviving firms that do not export and do not benefit from the reform.

For all other workers that got hired before the reform and at all time during the transition, wage premia can only increase as stated in equation (6). Using the latest notations, the premia increase by a factor $\left(\alpha'_{c,a,t}/\alpha'_{c,a-1,t-1}\right)^{\kappa\gamma} > 1$ when screening is increased from $\alpha'_{c,a-1,t-1}$ to $\alpha'_{c,a,t}$.

The workers already employed at the group C firms get a raise as their firm grows and screens more. Moreover, workers at the group C firms that export immediately get screened more and get a larger raise than what they would get without the reform. By contrast, workers at group B firms get screened less and get a lower raise (if at all anymore) than what they would get without the reform. Specifically, the workers at the group B firms that shrink keep the same wage until they quit or get fired (without screening).

The wage of new hires also depends on firm type. Because group B firms hire less than what they would do without the reform, they need not attract as many job candidate and they offer lower wage premia. Conversely, the group C firms that immediately export want to grow faster not to miss out too much on the new export opportunities: they need to attract more candidates and do so by offering wage premia that are higher than what they would offer without the reform.

So far, the above comparisons focus on the differences between the transition and the initial stead state. I characterize differences with the final steady state in the following two corollaries.

Corollary 2 to proposition 5 (Low productivity surviving firms)

There are firms with productivity $x < \underline{x}_\infty$ that survive for a period. Some of these firms have wages that are lower than any other wage in the future steady state.

Some surviving firms have a productivity level lower than any new entrants. These firms decided to enter and grow before the reform when the environment was not as competitive and their revenues were larger. Their size at the time of the reform is the advantage over new entering firms that allows them to survive despite a productivity lower than the new entry cutoff. These firms are the least productive: to survive, they are forced to pay low wages that no firms in the future steady state will ever offer. As exogenous exit shocks hit them, these unproductive firms gradually exit and are replaced by more productive firms

The firms that are old enough at t_0 to be in their stationary state act radically differently from the firms in the future steady state that have the same age. Because the reform is assumed to be unexpected, these firms could not smooth adjustments as future firms in the new steady state will do.

Corollary 3 to proposition 5 (Top wages and within-firm wage dispersion)
Consider the firms that were in their stationary state before t_0 and that belong to group C: (i) wage dispersion in their workforce increases, and (ii) these firms pay wages for some periods during the transition that are higher than what the firms with the same age and productivity would pay in the future steady state.

Firms that were close or at their stationary state just before the reform have not experienced growth for a long time. These old firms only hired workers to replace the ones who exogenously quit, implying that all workers were hired under the same condition of no firm growth: they paid the same wage to all their workers. With the reform, the group C firms start to grow again and offer large wage premia to speed the hiring process while saving on vacancy costs. The new hires are offered large wage premia and wage dispersion increases.

The above proposition and corollaries provide a good understanding of the mechanisms at play during the transition. In particular, low wages at barely surviving firms, wage cuts at the less productive firms that already pay little, and exceptionally high wage offers in the group of more productive firms that pay well contribute to more wage dispersion. However, the variety of firm experiences render the computation of aggregate variables intractable analytically and I turn to numerical solutions to obtain more results.

3 Calibration and aggregate results in the transition

In this section, micro-empirical evidence of the model central mechanisms are presented and used in a calibration exercise. The calibrated model allows for the study of aggregate variables during the transition path from autarky to trade.

3.1 Micro-evidence and calibration

The model introduces novel labor market mechanisms which need to be empirically established and quantified.

3.1.1 Taking wage offers and screening predictions to the data

Two important predictions of the model are investigated, the importance of screening at growing firms, and the relationship between firm growth and wage offers.

The model predicts that growing firms select better matches, but also separate with the least able of their existing workers: firms growing faster are predicted to separate with more workers. To test for and quantify this prediction, I estimate the following equation at the firm level:

$$\mathsf{sr}_{t,j} = \beta_{min} \min\{\mathsf{gr}_{t,j}, 0\} + \beta_{max} \max\{0, \mathsf{gr}_{t,j}\} + \alpha . X_{t,j} + \varepsilon_{t,j} \tag{33}$$

where j indexes firm, gr is the firm employment growth rate, and sr stands for the separation rate.[39] Several firm level characteristics[40] are included in $X_{j,t}$ to control for industry or location characteristics, and different sets are considered in robustness checks. Formally, the model predicts that the coefficient estimate of β_{max} is positive.

When considering wage offer predictions, an important caveat of the model is that it does not encompass worker observable characteristics, which are important features of wage data. I extend the model and assume that individuals, indexed by i, have observable characteristics that determine endowments in efficiency units of labor, e_i. Worker ability α is assumed to be independent from the labor endowments. The model is re-interpreted such that the labor variables, l, Δ, and V are expressed in efficiency units of labor. With these assumptions, firms are indifferent about how efficiency levels are distributed across workers and only care about the total number of units they hire and employ.[41] A worker endowed with e_i earns a wage $w_{r,i} = w_r\, e_i$ in the outside sector and a wage $w_{t,i} = w_t\, e_i$ in the differentiated sector.[42]

[39] gr is defined as the ratio of the difference in the number of workers between the beginning and the end of the period over the average number of workers in the period. sr is defined as the ratio of the number of workers who left the firm in the period over the same denominator.

[40] Workforce controls include the average wage and experience of incumbent workers and the share of hours worked by part time workers, female, senior staff, supervisors, clerical workers and blue collar workers. Other firm controls include revenues, capital intensity, materials, investment rate, imports and exports.

[41] I assume that the labor market is sufficiently large to avoid any situation in which firm behavior is constrained by the lack of availability of a particular number of labor units on the supply side.

[42] The outside option in the differentiate sector for a worker endowed with e_i is $w_{o,i} = w_o\, e_i$.

The analysis of the theory section carries through and the wage of worker i offered by firm j at t depends on observable characteristics, $e_{t,i}$, the hiring rate, $\left(\Delta/\check{l}\right)_{t,j}$, and the screening level, $\alpha_{t,j}$, of the firm:

$$w_{t,i,j} - w_{o,t} = e_{t,i}\frac{(1-\xi)}{\gamma}\frac{c_A}{c_w \alpha_{min}^{\kappa\gamma}}\left(\frac{\Delta}{\check{l}}\right)_{t,j}^{\gamma-1}(\alpha_{t,j})^{\kappa\gamma} \tag{34}$$

I take this equation in logs to the data in two steps by estimating

$$\ln(w_{t,i,j} - w_{o,t}) = \beta_0 + \beta_1 . \ln e_{t,i} + \varepsilon_{t,i} \tag{35}$$

$$\ln \overline{w}_{\text{res.},t,j}^{hires} = \ln\left(\frac{(1-\xi)}{\gamma}\frac{c_A}{c_w \alpha_{min}^{\kappa\gamma}}\right) + (\gamma-1) . \ln\left(\Delta/\check{l}\right)_{t,j} + \kappa\gamma . \ln\left(\alpha_{t,j}\right) + \varepsilon_{t,j} \tag{36}$$

with $\ln \overline{w}_{\text{res.},t,j}^{hires} = \frac{1}{\Delta_{t,j}}\sum_{i\in j}\widehat{\varepsilon}_{t,i}$ and where $\varepsilon_{t,i}$ and $\varepsilon_{t,j}$ are measurement errors. For equation (35), I run a Mincer-type regression on the log demeaned wage of all new hires at the time of hiring to partial out the observable worker characteristics[43] and recover the wage residuals. This is done at the individual level, separately for every year. Then, I compute the average residual wage of new hires at every hiring firms in every year, $\ln \overline{w}_{\text{res.},t,j}^{hires}$ and regress it on firm hiring rates. The model predicts that the coefficient estimate of the hiring rate should be positive since $\gamma > 1$.

There are two challenges in estimating equation (36). First, the screening level $\alpha_{j,t}$ is not observed. I include firm fixed effects in the regression, as well as firm time-varying characteristics as proxies for the screening level. In particular, the average wage of incumbent workers is included because the model predicts that it should be directly related to the current screening level. Second, firm level shocks to the recruiting technology or shocks to the labor supply captured by changes in c_A, c_w or α_{min}, can generate an omitted variable bias. Specifically, one should expect a better recruiting technology or positive local shocks to the labor supply to simultaneously raise firm growth and lower the wage of new hires, implying a downward bias in the estimation of γ.

To address these endogeneity concerns, I restrict the sample to exporters and instrument firm growth with the foreign demand shock developed in Hummels et

[43] Controls include a dummy for living in the Paris area, a part-time status dummy, a gender dummy, a polynomial of order four in labor market experience, and occupation-industry dummies

al. (2014):[22] the average changes in foreign partners' demand. I use the weighted average change in the imports of foreign countries from the rest of the world but France. The weights of every product-destination are firm specific: they are computed as the share of the product-destination export values in total exports in the first year of observed exports.

3.1.2 Micro data

For the parts of the calibration requiring estimation, I use a combination of datasets from France covering the period 1995-2007. The combined dataset has a number of characteristics that make it suitable for the calibration. First, it contains both workers and firms characteristics. Second, it allows for the characterization of gross worker flows at the firm level. Third, France labor regulation is average: for the three employment protection indexes published by the OECD for high and middle income countries, the values for France are within one standard deviation of the average values of middle-income countries.[44] Hence, French data are a reasonable candidate to calibrate labor market parameters.

I obtain worker information from the matched employer-employee DADS.[45] This dataset is based on mandatory employer reports of the earnings of each employee. While it covers the full universe of the private sector, I only use information on the individuals that worked once for a manufacturing firm. The data allows me to identify the number of hires and separated workers at the firm level and their characteristics. The DADS has information on workers' gender, age, and location as well as information on jobs' occupation type, earnings and full or part time status.

I match the DADS with the census of manufacturing firms of 20 employees and over (henceforth, the EAE).[46] Hence, I complement information on workforce with firm characteristics on value added, profits, capital, and investment. The firm level quantities and values of imports and exports by country and product category that

[44] I choose to compare France with middle income countries because economies that implemented trade reforms typically had middle income.
Link to data: http://www.oecd.org/employment/emp/oecdindicatorsofemploymentprotection.htm

[45] The *Déclarations Annuelles de Données Sociales* (DADS) dataset is put together by the French National Institute of Statistics and Economics Studies (INSEE)

[46] The *Enquête Annuelle d'Entreprise* is also collected by the INSEE.

Table 1: The relationship between separation rates and firm growth

Dependent variable: the separation rate	(1.A) Basic	(1.C) FE, Weights & controls	(1.E) Revenue growth	(1.F) Smoothed growth
max(0,growth rate)	0.394***	0.125***	0.033***	0.081***
	(0.008)	(0.009)	(0.004)	(0.007)
min(growth rate,0)	-0.992***	-0.839***	-0.072***	-0.563***
	(0.006)	(0.007)	(0.008)	(0.019)
Firm effects	NO	YES	YES	YES
Controls	NO	YES	YES	YES
Weights	NO	# of jobs	# of jobs	# of jobs
Observations	209,264	187,896	189,132	187,896
R-squared	0.179	0.691	0.563	0.588

Details and coefficient estimates of the controls listed in footnote 40 are in appendix. Firm-level clustered standard errors are in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

are needed to construct the instrument are obtained from customs data.[47]

3.1.3 The variation of wages and screening with firm growth in the data

Estimation results of equation (33) presented in table 1 shows that the separation rate is increasing in the employment growth rate, as predicted in the model.[48] This is not a unique feature of French data, as the same pattern was recorded for the U.S in graphs by Davis et al. (2012). [10] The correlation between firm growth and separation rates is robust to the inclusion of firm fixed effects, weights and various controls (column 1.C), using revenue growth (column 1.E) or a three-year average of employment growth (column 1.F) as alternative measures of firm growth.

Estimation results of equation (36) presented in table 2 show that wage offers to new hires increase significantly with firm growth rates. The OLS estimate of the hiring rate coefficient in column 2.A is 0.19. As expected, the new estimate is larger as expected once the hiring rate is instrumented (column 2.B). The first

[47] The *Données import/export du commerce extérieur* are collected by the Direction Générale des douanes et des droits indirects.

[48] The result may come as a surprise given that French labor market is commonly viewed as rather rigid. However, the OECD indexes suggest that it is not extremely rigid. Moreover, despite strong labor protection for employees under open-ended contracts, the existence of short term contracts allows for some additional flexibility from the firm perspectives. Firms can hire workers on fixed term contracts for up to three years and many separations include the end of these.

Table 2: Firm level estimation of the elasticity of wage offers to firm growth

Dependent variable: new hires avg. residual wage	2.A OLS-FE	2.B IV-FE	2.C additional controls	2.D Single estb.
log hiring rate $\ln\left(\Delta/l\right)_{j,t}$	0.019***	0.083*	0.099*	0.09
	(0.002)	(0.043)	(0.057)	(0.089)
incumbents' avg. wage	0.540***	0.511***	0.527***	0.498***
	(0.017)	(0.028)	(0.028)	(0.042)
Observations	96588	96588	89376	70122
Validity tests of the first stage:				
– F-test of excluded IV		0.0012	0.0092	0.17
– Weak identification (max. IV size)		10%	10%	20%

All specifications include additional controls, year and firm fixed effects, and are weighted by the number of hires. Details and additional coefficient estimates are in appendix. Firm-level clustered standard error are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

stage is valid as shown by the test results. The corresponding estimated value of γ is 1.10 and corresponds to modestly convex adjustment costs. The estimates are robust to restricting the sample to single establishment firms (columns 2.C and 2.D). Altogether, these estimates are somewhat larger but close to the 5% estimate in Schmieder (2013).[28] The difference in results could come from the fact that Schmieder (2013) focuses on small new firms rather than on exporters which are more relevant for my analysis.

3.1.4 Model calibration

For the numerical exercise, I assume that the reduction in trade costs is such that openness in the future steady state matches French data characteristics:[49] specifically I assume that 10% of firms export and that they export 30% of their sales.

The model has a number of standard features that have been extensively studied in the literature. Therefore and whenever readily available, I use standard estimates and calibration methods from the literature.I choose periods to correspond to years. I follow Ghironi and Melitz (2005)[16] and choose the discount rate to be $\beta = 0.95$, the inter-temporal elasticity of substitution $IES = 2$, and the elasticity of substitution between varieties $\sigma = 4$. The coefficient of diminishing returns to labor $\rho = 0.7$ is set to obtain a wage to valued added ratio of 50%.

I follow the methodology of Head, Mayer and Thoenig (2014)[18] for the calibra-

[49] Trade openness in 1995-2007 France is slightly below 50%, close to the world average of 50%.

tion of the firm productivity distribution and the levels of trade costs in the terminal steady state. Assuming a Pareto distribution, I recover the shape parameter from the right tail of the distribution of firm value added. The exogenous exit rate of firms $\delta_0 = 2.5\%$ is based on the observed exit rate in the EAE data.[50]

Normalization for population, productivity, ability and prices allows me to choose the values of population, the coefficient of the technology cost (c_S), the entry and operation costs and to set the outside option as the numéraire.[51] The triplet (κ, ρ, ψ) is over-identified[52] and I choose $\psi = 7$ for clarity of exposition.

The remaining parameters pertain to the labor market. The quit rate $s_S = 0.19$ is measured from the data by assuming it corresponds to the separation-to-employment ratio when there is no firm growth. The search cost coefficient $c_u = 4$ is set to obtain a sectoral unemployment rate of 9%. I choose the share of vacancy costs to wage offers, $\xi = 10\%$, to match the magnitude of recruitment costs reported in Manning (2011).[24] The convexity of adjustment was previously estimated at $\gamma = 1.1$. The shape parameter of the ability distribution, κ, and the coefficient of the adjustment cost function, c_A, are set respectively at 1.05 and 5 to match the elasticity of separation rates to firm growth (estimated at 13% in the data) and the elasticity of firm average wages to employment (estimated at 10% in the data).

3.2 The transition path following a trade liberalization reform

Equipped with the calibrated model, I compute the predicted numerical responses of aggregates to the once and for all opening of the economy. Details of computations are in appendix but the algorithm used to solve for the equilibrium path essentially relies on searching for a fixed point by iteration, updating $(\phi_t, \mathscr{P}_t)_{t=1..T}$

[50] In the EAE data, 7% of firms exit the census every year. That rate includes administrative restructuring and changes in the the firm administrative identifiers. Therefore, I adjust this rate according to the findings of Hethey-Maier and Schmieder (2013)[28] showing that only one third of identification code terminations corresponds to real firm exits.

[51] The entry cost f_E has no other impact than determining the unobserved fraction of successful entrants. I arbitrarily choose $f_E = 1.85$. The operation costs, $f_d = 0.1$, and the coefficient of the technology costs, $c_S = 0.05$, are chosen for convenience of exposition. Population is set to one.

[52] Intuitively, the elasticity of the cost function ψ, the strength of complementarities and the dispersion of abilities have offsetting effects on the net returns to screening: greater net returns are obtained from raising the cutoff when (i) the cost-elasticity is lower, or (ii) ability dispersion is greater or (iii) complementarities are stronger

guesses with the **FE-ZCP** and **LMC** conditions.

3.2.1 Consumption, price, firm creation and trade results

Figure 3 shows that production, and thus consumption,[53] increase and overshoot their long run value as in Ghironi and Melitz (2005) [16] or Alessandria and Choi (2014). [2] The presence of sunk costs introduces inertia and allows unproductive firms that would not successfully enter in the new steady state to stay active for a period (Corollary 2 to proposition 5). These firms' production more than compensates the sluggish expansion of exporters and total production overshoots.

At the beginning of the transition, the shift of labor resources towards production and the saturated product market raise the entry cost relative to gains and deter firm creation. Rising costs are reflected by a drop in the price index which falls below its long run value. Lower firm creation translates into a temporary drop in the mass of entrants.

In the second phase of the transition, production falls to levels below the future steady state level but gradually recovers to reach it from below. This results from misallocation of labor in the production sector: there are too many workers still employed at old unproductive firms and not enough at new more productive ones. Among older firms that existed before the reform, the very unproductive ones that will be replaced by better firms are still around retaining some workers. In addition, lower creation of more productive firms at the beginning of the transition creates a shortage of mature productive firms during the second phase of the transition when the firms born at the beginning of the transition reach full capacity.

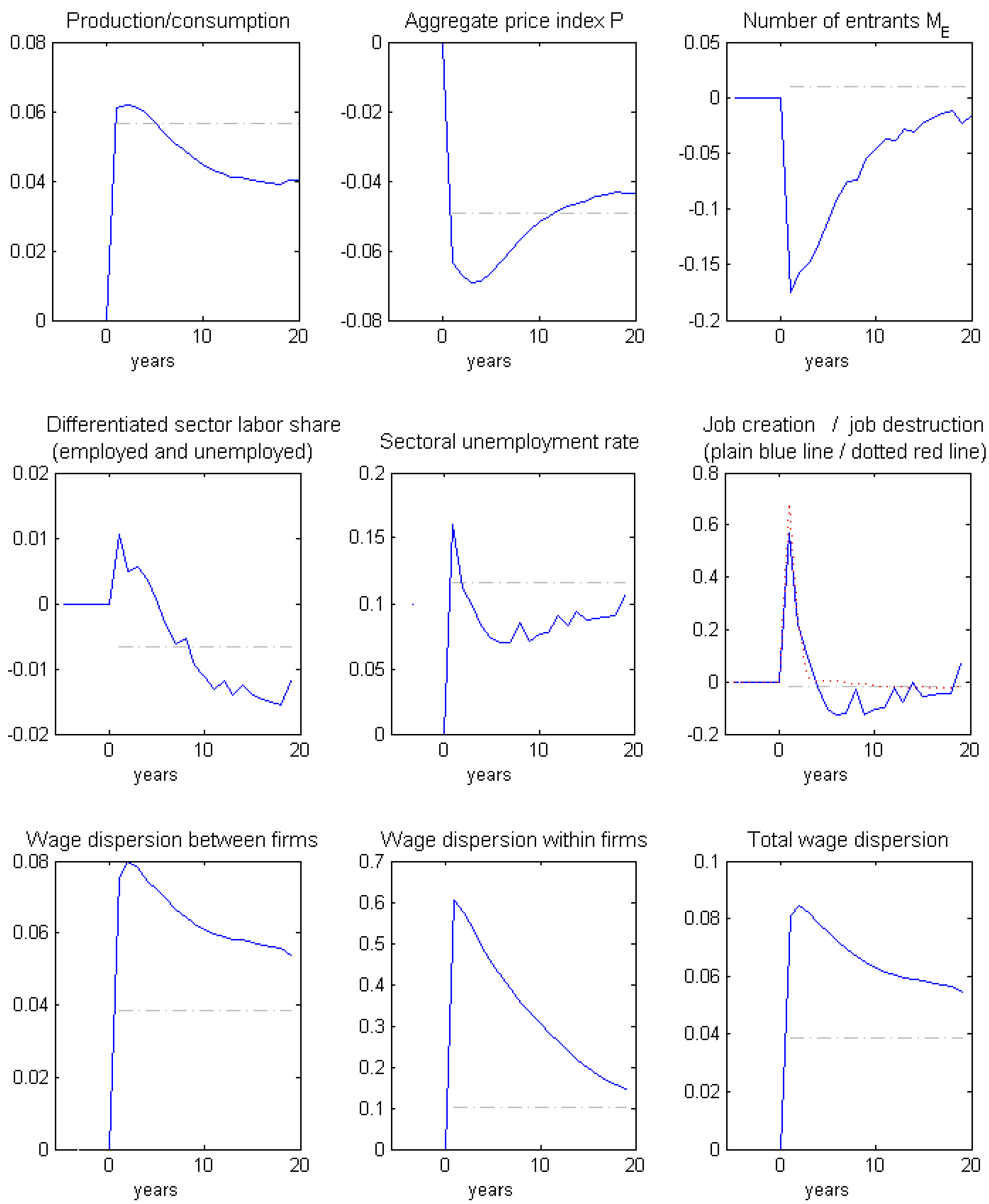
3.2.2 Labor allocation and the evolution of inequality

The long run increase in inequality, unemployment and service sector employment are consistent with propositions 3 and 4. Short term variations are as follows.

Unemployment overshoots its long run value. Variations in unemployment are essentially determined by the length of vacancy queues, which are the result of wage offer premia. At the beginning of the transition, unemployment increases above its

[53] Because of the absence of a storage technology and the assumption of balanced trade, consumption and production are equal at all times.

Figure 3: Evolution of selected aggregates in log-deviation from the autarky level.



Notes: Inequality is measured by coefficients of variation. The dashed gray lines represent the future steady state values towards which the economy converges.

new steady state level because current and future exporters offer large premia to attract better workers fast and speed their expansion.

Employment in the differentiated sector first rises, then drops substantially and reach the new steady state level from below. At the beginning of the transition, the expansion of exporter employment more than compensates the decline of domestic firms. In the meantime, service employment for firm creation drops. In the second phase of the transition, employment in the differentiated sector drops as firm exits are only partially compensated by the employment growth of the too few new firms born after the reform. Meanwhile, the high number of surviving unproductive firms still necessitate more than usual labor services for their operation cost (f_d).

Figure 3 also shows that inequality increases and overshoots its long run value. Quantitatively, the coefficient of variation increases by about 8% in the first 3 years, but a fourth of that increase gets undone in the following decade. Inequality is mainly driven by the dispersion in firm average wages. At the top of the wage distribution, the high-paying current and future exporters raise their screening levels and offer wage premia during their expansion. In the future steady state, their workers hired with high wage premia during the transition will have retired. They will be replaced by new workers hired with lower wage premia. At the bottom of the distribution, workers at unproductive firms accept wage cuts to delay their employer exit and preserve their jobs. These workers will eventually separate from these firms and they will be replaced by workers that get jobs at better firms with better wage.

There is also a large temporary increase in within-firm inequality. This increase results from the difference between the wage of incumbent workers and the wage of new hires that are recruited with a premium at expanding firms. However the contribution of within inequality to total inequality is small because within-firm inequality is very small relative to between-firm inequality.

4 Conclusion

If the literature shows that trade liberalization causes inequality to rise, little was known about the dynamics of the inequality response across firms and workers. This paper contributes to the literature by developing a trade model of firm and worker heterogeneity with endogenous firm dynamics.

Equivalence steady state results with canonical models show that the model is macro-relevant: long run welfare and inequality increase with openness. Following a trade reform, I obtain analytical results about firm and worker transition paths. The model is well micro-founded as shown by micro-empirical evidence from a matched employer-employee data from France.

The calibrated model predicts the overshooting of inequality and is thus able to replicate a macro-data pattern. It highlights the role of the reallocation process in shaping the wage distribution and its evolution. Inequality increases faster in the model than in the data. This could come from the fact that liberalization reforms tend to be gradual, unlike what is assumed in the paper. Alternatively, the excessively fast adjustment of inequality may result from the absence of other adjustment frictions. For example, destination specific frictions would slow the reallocation of workers. These observations highlight a path for future research.

References

[1] George A. Akerlof, William R. Dickens, and George L. Perry. The Macroeconomics of Low Inflation. Brookings Papers on Economic Activity, 27(1):1–76, 1996.

[2] George Alessandria and Horag Choi. Establishment heterogeneity, exporter dynamics, and the effects of trade liberalization. Journal of International Economics, 94(2):207–223, 2014.

[3] Costas Arkolakis, Arnaud Costinot, and Andres Rodriguez-Clare. New Trade Models, Same Old Gains? American Economic Review, 102(1):94–130, February 2012.

[4] John M. Barron, Mark C. Berger, and Dan A. Black. On-the-Job Training. Number ojt in Books from Upjohn Press. W.E. Upjohn Institute for Employment Research, January 1997.

[5] Marc Blatter, Samuel Muehlemann, and Samuel Schenker. The costs of hiring skilled workers. European Economic Review, 56(1):20–35, 2012.

[6] David Card, Jörg Heining, and Patrick Kline. Workplace Heterogeneity and the Rise of West German Wage Inequality. The Quarterly Journal of

Economics, 128(3):967–1015, 2013.

[7] Thomas Chaney. Distorted Gravity: The Intensive and Extensive Margins of International Trade. American Economic Review, 98(4):1707–21, 2008.

[8] Kerem Cosar. Adjusting to trade liberalization: Reallocation and labor market policies. University of Chicago Booth School of Business, unpublished manuscript, 2010.

[9] Eliav Danziger. Skill acquisition and the dynamics of trade-induced inequality, 2014.

[10] Steven J Davis, R Jason Faberman, and John Haltiwanger. Labor market flows in the cross section and over time. Journal of Monetary Economics, 59(1):1–18, 2012.

[11] Rafael Dix-Carneiro. Trade Liberalization and Labor Market Dynamics. Working Papers 1273, Princeton University, Department of Economics, Center for Economic Policy Studies., November 2010.

[12] Arindrajit Dube, Eric Freeman, and Michael Reich. Employee Replacement Costs. Institute for Research on Labor and Employment, Working Paper Series qt7kc29981, Institute of Industrial Relations, UC Berkeley, March 2010.

[13] Pablo D. Fajgelbaum. Labor Market Frictions, Firm Growth, and International Trade. NBER Working Papers 19492, National Bureau of Economic Research, Inc, October 2013.

[14] Gabriel Felbermayr, Giammario Impullitti, and Julien Prat. Firm Dynamics and Residual Inequality in Open Economies. IZA Discussion Papers 7960, Institute for the Study of Labor (IZA), February 2014.

[15] Judith A. Frias, David S. Kaplan, and Eric Verhoogen. Exports and wage premia: Evidence from mexican employer-employee data. W.p, 2009.

[16] Fabio Ghironi and Marc J. Melitz. International Trade and Macroeconomic Dynamics with Heterogeneous Firms. The Quarterly Journal of Economics, 120(3):865–915, August 2005.

[17] Pinelopi Koujianou Goldberg and Nina Pavcnik. Distributional Effects of Globalization in Developing Countries. Journal of Economic Literature, 45(1):39–82, March 2007.

[18] Keith Head, Thierry Mayer, and Mathias Thoenig. Welfare and Trade without

Pareto. American Economic Review, 104(5):310–16, May 2014.

[19] Elhanan Helpman, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen J. Redding. Trade and inequality: From theory to estimation. NBER Working Papers 17991, National Bureau of Economic Research, Inc, April 2012.

[20] Elhanan Helpman, Oleg Itskhoki, and Stephen Redding. Inequality and unemployment in a global economy. Econometrica, 78(4):1239–1283, 07 2010.

[21] Christian Holzner and Mario Larch. Capacity Constraining Labor Market Frictions in a Global Economy. Technical report, 2011.

[22] David Hummels, Rasmus Jørgensen, Jakob Munch, and Chong Xiang. The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data. American Economic Review, 104(6):1597–1629, June 2014.

[23] Leo Kaas and Philipp Kircher. Efficient Firm Dynamics in a Frictional Labor Market. American Economic Review, 105(10):3030–60, October 2015.

[24] Alan Manning. Imperfect Competition in the Labor Market, volume 4 of Handbook of Labor Economics, chapter 11, pages 973–1041. Elsevier, 2011.

[25] Marc J. Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica, 71(6):1695–1725, November 2003.

[26] Monika Merz and Eran Yashiv. Labor and the Market Value of the Firm. American Economic Review, 97(4):1419–1431, September 2007.

[27] Julian Emami Namini and Ricardo A. López. Factor price overshooting with trade liberalization: theory and evidence. Scottish Journal of Political Economy, 60(2):139–181, 05 2013.

[28] Johannes Schmieder. What causes wage dispersion? evidence from new firms. Technical report, 2013.

[29] Romain Wacziarg and Jessica Seddon Wallack. Trade liberalization and intersectoral labor movements. Journal of International Economics, 64(2):411–439, December 2004.