PRODUCTIVITY DISPERSION, BETWEEN-FIRM COMPETITION, AND THE LABOR SHARE

ÉMILIEN GOuin-BONENFANT
Department of Economics, Columbia University

I study the effect of labor market imperfections on the labor share in a tractable model that emphasizes the interaction between productivity dispersion and firm competition for workers. I calibrate the model using administrative data covering the universe of firms in Canada from 2000 to 2015. As in the data, most firms have a high labor share, yet the aggregate labor share is low due to the disproportionate effect of a small fraction of large, highly productive firms. I find that a rise in the dispersion of firm productivity causes the aggregate labor share to decline in favor of firm profits. The mechanism is that productivity dispersion effectively shields high-productivity firms from wage competition. Regression evidence from cross-country and cross-industry data supports both the model prediction and mechanism.

KEYWORDS: Productivity dispersion, Firm dynamics, Search frictions, Labor share.

1. INTRODUCTION

For most of the twentieth century, worker compensation in the U.S. grew one-for-one with labor productivity. This observation has been enshrined as one of the stylized facts of economic growth (Kaldor, 1961). However, starting around 1990, worker compensation has stagnated relative to labor productivity, leading to a decline of the labor share (Elsby, Hobijn, and Şahin, 2013; Karabarbounis and Neiman, 2014). Over that same period, there has been a sustained rise in the dispersion of firm productivity, where the labor productivity gap between the best firms and the rest has widened (Kehrig, 2015; Barth, Bryson, Davis, and Freeman, 2016).

In this paper, I study the effect of labor market imperfections on the labor share in a model that emphasizes the interaction between productivity dispersion and firm competition for workers. First, I build a tractable model of firm dynamics with search frictions and wage posting in the labor market. I use the model to generate theoretical predictions regarding the link between the distribution of firm productivity and the aggregate labor share. Second, I calibrate and validate the model using administrative data covering the universe of firms in Canada over the 2000–2015 period. Third, I use the model to simulate the effect of a rise in productivity dispersion and contrast its predictions with evidence from cross-industry and cross-country data.

I have two main findings. First, I find that labor market imperfections have large and heterogeneous effects on firm-level labor shares. Through the lens of the model, labor market imperfections depress the aggregate labor share by 15 percentage points (compared to a frictionless model) yet they raise the average firm-level labor share by 12 percentage points. This asymmetric effect is due to the fact that value-added is concentrated within large, highly productive firms that exert significant monopsony power (i.e., they pay wages below the marginal product of labor).

Second, I establish theoretically that a rise in productivity dispersion causes the aggregate labor share to decline in favor of firm profits. The amount of competition that firms face in the model is effectively determined by how close they are to their competitors in terms of

Émilien Gouin-Bonenfant: eg3041@columbia.edu
This is a revised version of the first chapter of my dissertation at UCSD. I wish to thank my advisors Jim Hamilton, David Lagakos, Tommaso Porzio, Natalia Ramondo, Allan Timmermann, and Alexis Akira Toda. I am also grateful to Isaac Sorkin for discussing my paper as well many others for helpful comments and suggestions.
productivity. Rising productivity dispersion thus shields high-productivity firms from wage competition. Quantitatively, I find that a rise in productivity dispersion of the same magnitude as what the U.S. economy has experienced over the 2000–2015 period causes a 1.3 percentage point decline in the aggregate labor share. Interestingly, the effect of a rise in productivity dispersion works through a reallocation of value-added towards firms with a low (and declining) labor share, rather than a broad-based decline of firm-level labor shares.

**Overview of the paper.** In Section 2, I present the model and derive a number of theoretical predictions. The model embeds a labor market with search frictions, wage posting, and on-the-job search (as in Burdett and Mortensen, 1998) in a continuous-time firm dynamics model where firms face productivity shocks (as in Coles and Mortensen, 2016). In the model, firms solve a dynamic problem and choose how much capital to rent, what wage to offer, and when to exit. In equilibrium, high-productivity firms offer high wages in order to poach workers from lower-paying firms and therefore grow faster. Relative to existing wage-posting models, my environment includes capital as a factor of production as well as endogenous firm entry and exit. These ingredients are key: introducing capital allows the model to generate an empirically reasonable level for the labor share and free entry ensures that ex-ante rents are exhausted.

I obtain an analytical solution for the stationary equilibrium of the model, which allows me to generate a number of new theoretical predictions. At the micro level, I show that firm-level labor shares are (asymptotically) decreasing in firm productivity. The reason is that the pass-through of productivity to wages is proportional to the (employment-weighted) density of firms. Firms in the right tail of the productivity distribution have few direct competitors and are therefore able to recruit and retain workers despite paying wages below the marginal product of labor. In contrast, the least productive firms pay wages that are above their marginal product, and therefore have a high labor share. For these firms, the option value of retaining their workers compensates for the negative flow profits.

The main theoretical contribution of the paper is to show that the aggregate labor share is decreasing in the level of productivity dispersion (i.e., the “thickness” of the right tail of the productivity distribution). To establish this result, I focus on a special case where the productivity distribution is Pareto and the discount rate is zero. I show that, all else equal, the aggregate labor share is decreasing in productivity dispersion. The intuition is that when firms in the right tail of the productivity distribution become “further apart” in terms of productivity, they effectively face less competition and thus earn higher monopsony profits. I also show that a key determinant of the sensitivity of the labor share to a change in productivity dispersion is the persistence of productivity differentials across firms. A more persistent productivity process implies a larger labor share decline in response to a rise in productivity dispersion. Intuitively, what matters is not only how large productivity differentials across firms are, but also how persistent they are.

In Section 3, I calibrate and validate the model using administrative data covering the universe of corporations in Canada over the 2000–2015 period. The dataset is ideal as it allows me to compute labor productivity, labor share, capital stock, and employment at the firm level. I calibrate the model by targeting moments related to the firm productivity process as well as the pace of job reallocation between firms. Despite being very parsimonious, the model matches a number of non-targeted (within-industry) moments. For example, the aggregate labor share is 0.62 (0.65 in the data) while the average (unweighted) firm labor share is 0.89 (0.88 in the data). The reason why the average labor share is so high is that the bottom 3 deciles of firms

---

1The precise statement is that there exists a productivity threshold after which the labor share of a firm is decreasing in its productivity.
in terms of labor productivity have labor share above one (bottom two deciles in the data). The model also predicts that value-added is concentrated within high-productivity, low-labor-share firms: the top decile of firms in terms of labor productivity account for 55 percent of output (42 percent in the data) and have a labor share of 0.53 (0.44 in the data). Finally, I assess the ability of the model to match “panel moments” such as the pass-through of labor productivity to wages and find that the model provides a satisfactory fit.

In Section 4.1, I use the calibrated model to simulate the effect of a rise in productivity dispersion. To motivate the model experiment, I show that, in recent decades, there has been a divergence between the U.S. and Canadian economies. Both productivity dispersion and the labor share have been stable in Canada, while the U.S. has experienced a rise in productivity dispersion coupled with a decline of the labor share.

The model experiment consists of simulating the effect of a rise in productivity dispersion of the same magnitude as what the U.S. experienced over the 2000–2015 period. The model predicts a labor share decline of 1.3 percentage points. Using a shift-share decomposition, I show that the decline of the labor share in the model is mostly driven by a reallocation of value-added shares between firms, rather than a broad-based decline of firm-level labor shares. High-productivity firms experience an increase in their value-added share as well as a decrease in their labor share. In contrast, low-productivity firms experience a decrease in their value-added share combined with an increase in their labor share. Through the lens of the model, roughly one quarter of the divergence between the Canadian and U.S. labor shares over the 2000–2015 period can be explained by differential trends in productivity dispersion.

The micro predictions of the model are in line with what has been documented by Autor, Dorn, Katz, Patterson, and Van Reenen (2020) and Kehrig and Vincent (2021) using U.S. firm-level data. In particular, the model predicts that in response to an increase in productivity dispersion: (1) the aggregate labor share declines, (2) the decline is driven by a reallocation of value-added towards firms with a low (and declining) labor share, and (3) sales concentration increases.

In Section 4.2, I compare the predictions of the model with regression evidence from cross-country and industry data. First, I use industry data from 69 Canadian 3-digit industries over the 2000–2015 period. I estimate cross-sectional, fixed effects, and long-differences regression models. In all cases, I find that industries with high (rising) productivity dispersion have a low (declining) labor share. The results are robust to controlling for other factors such concentration (as in Barkai, 2020, Autor et al., 2020) and the capital-output ratio (as in Karabarbounis and Neiman, 2014). I also test the model-implied mechanism by estimating the effect of productivity dispersion along the firm productivity distribution. As predicted by the model, a productivity dispersion shock has an asymmetric effect on firm labor shares and leads to a reallocation of value-added towards high-productivity firms. Second, I collect harmonized data on productivity dispersion and the labor share for 7 OECD countries over the 2001-2011 period and find quantitatively similar results.

Finally, in Section 5, I consider a number of model extensions (including the case with endogenous vacancy creation) in order to assess how sensitive the results of the model experiment are to different modeling choices. Overall, I find that the baseline model is somewhat conservative, in the sense that most of the model extensions considered imply a larger labor share decline in response to a rise in productivity dispersion.

Related literature. The main contribution of this paper is to link the rise in productivity dispersion to the decline of the labor share, both theoretically and empirically. A related find-

2Several studies document a decline of the labor share in the US and other countries (see, e.g., Elsby, Hobijn, and Şahin, 2013, Karabarbounis and Neiman, 2014). A less appreciated fact is that productivity dispersion has increased
ing emerges in Hartman-Glaser, Lustig, and Xiaolan (2019), who study the link between the volatility of sales and the capital share in a firm dynamics model where risk-neutral firms insure risk-averse workers. Through the lens of their model, they find that an increase in firm-level risk generates an increase in the aggregate capital share. My mechanism differs in that productivity dispersion reduces the effective amount of competition faced by high-productivity firms, while they emphasize the insurance contract between workers and shareholders.

My findings complement a growing literature that studies the importance of firm heterogeneity in shaping the decline of the U.S. labor share. Kehrig and Vincent (2021) focus on the manufacturing sector over the 1967–2012 period. They find that the decline of the labor share was driven by a reallocation of value-added towards firms with a low (and declining) labor share, but that there was limited reallocation of inputs (labor and capital). Autor et al. (2020) use data covering most sectors of the U.S. economy over the 1982–2012 period. They also find that the labor share decline was mostly driven by reallocation of market shares between firms, as opposed to a broad-based decline of labor shares within firms. In addition, they also show that industries that saw the largest increases in sales concentration also saw the largest declines in their labor shares. I contribute to this literature by providing (i) new evidence on the relationship between labor share and labor productivity at the firm-level, (ii) a theory that can explain the large heterogeneity in labor shares across firms, and (iii) a new mechanism that generates a “reallocation-driven” labor share decline.

My paper contributes to the literature on the labor share by studying the aggregate effects of search frictions. Jarosch, Nimczik, and Sorkin (2019) study a “granular” search model, where wages are determined by Nash bargaining. The main difference with my paper is that they study market power as arising due to firm size, while in my model there is no market power in the traditional sense (firms are atomistic).3 Instead, I focus on the rent-sharing problem between workers and firms, and in particular how productivity dispersion affects the optimal wage-setting decision of firms.

Relative to the frictional monopsony literature—which studies how monopsony power arises in the presence of search frictions and wage posting—my main contribution is to establish a theoretical link between productivity dispersion and the aggregate labor share and to quantify its importance.4 The most closely related papers are Bontemps, Robin, and Van den Berg (2000) and Cahuc, Postel-Vinay, and Robin (2006). Both papers structurally estimate search models using French administrative data on wages and firm productivity. Cahuc et al. (2006) quantifies the contributions of bargaining power and between-firm competition in raising wages above the value of unemployment. They find that between-firm competition is the most important factor. Bontemps et al. (2000) use a wage-posting model and find that high-productivity firms exert more monopsony power, which is closely related to my finding that high-productivity firms have a lower labor share. Compared to these papers, who focus on the distribution of wages, over time. Evidence for the U.S. include Kehrig (2015), Barth, Bryson, Davis, and Freeman (2016) and Decker, Haltiwanger, Jarmin, and Miranda (2018). Evidence for OECD countries include Andrews, Criscuolo, Gal et al. (2016) and Berlingieri, Blanchenay, and Criscuolo (2017).

3Several recent papers study the aggregate implications of market power in frictionless models. For instance, see Barkai (2020), Eggertsson, Robbins, and Wold (2021), De Loecker, Eeckhout, and Unger (2020), and Edmond, Midrigan, and Xu (2018) for product market power and Azar, Marinescu, and Steinbaum (2020), Benmelech, Bergman, and Kim (2020), Berger, Herkenhoff, and Mongey (2019), and Brooks, Kaboski, Li, and Qian (2021) for labor market power.

4Recent applications of the Burdett and Mortensen (1998) model to study the wage-setting behavior of firms include: Meghir, Narita, and Robin (2015), who study the role of informal labor markets; Engbom and Moser (2021), who study the effect the minimum wage on inequality in Brazil; Heise and Porzio (2021) who study the role of firms and location preferences in shaping spatial wage gaps; and Bilal and Lhuillier (2021), who study the effect of outsourcing.
my paper focuses on the aggregate labor share, and in particular how productivity dispersion affects the aggregate profit share.


In terms of the environment, my model differs from Coles-Mortensen along three main dimensions. First, I add capital as a factor of production. Payments to capital affect the share of value-added that accrues to workers. Second, I model the entry and exit decision of firms, while Coles-Mortensen assumes that firm productivity is high enough so that firms always make positive profits and therefore always wish to enter and never wish to exit. Free entry ensures that ex-ante rents are exhausted (i.e., the marginal entrant makes zero profits in expectation), thereby limiting the amount of profits in equilibrium. Moreover, endogenous exit allows the model to rationalize the existence of temporary negative profits (an important feature of the data). Finally, Coles-Mortensen model the hiring process using a hiring cost function. Instead, I assume “balanced matching” as in Burdett and Vishwanath (1988) (i.e., exogenous meeting rates that are proportional to firm size), which delivers analytical tractability and allows me to obtain my main theoretical results.\(^5\) Finally, in Section 5.3, I relax the balanced matching assumption by considering a model extension with endogenous vacancy-posting.

2. THEORY

I now present the baseline model. I start by describing the environment and assumptions, paying particular attention to dimensions in which it differs from Burdett and Mortensen (1998) and Coles and Mortensen (2016). The key theoretical results are in Section 2.3, where I obtain model predictions regarding the effect of productivity dispersion on the aggregate labor share.

2.1. Environment

The economy is populated by an endogenous measure \( F \) of heterogeneous firms and a unit measure of identical, infinitely-lived workers. Firms and workers are risk neutral and discount the future using an exogenous discount rate \( r > 0 \). Time is continuous.

Technology. Firms compete for workers by posting wage contracts \( w \) in a frictional labor market and rent capital in a perfectly competitive market at user cost \( R \equiv r + d \), where \( d \geq 0 \) is the depreciation rate of capital. Firms can change the wage policy at any time and can not commit to future wages. I assume that firms are required to pay all of their workers the same wage. Workers can be either employed or unemployed and are available to receive job offers in both states. The flow value of unemployment is exogenous and given by \( b > 0 \).

\(^5\)In particular, I show that the stationary equilibrium of the model is the solution to a system of quadratic equations (Proposition 3). The resulting tractability of the model allows me to establish my main theoretical result (Proposition 6). In terms of new theoretical results, I also provide a sufficient condition for the existence of a stationary equilibrium (Assumption 2) and a characterization of the Pareto exponent of the firm-size distribution (Appendix Proposition 9). I use theoretical results on power laws (i.e., Toda, 2014; Beare, Seo, and Toda, 2021; Beare and Toda, 2022) and the “Pareto extrapolation” solution method (developed in Gouin-Bonenfant and Toda, 2022) to solve the firm size distribution. The emergence of fat-tailed size distribution in my model echoes findings in Luttmer (2007) and Luttmer (2011), who show that homogeneous firm dynamics models naturally give rise to fat-tailed distributions.
Firm output $Y$ is given by a constant return to scale Cobb-Douglas production function

$$Y = zK^\alpha N^{1-\alpha},$$

where $z$ is productivity, $\alpha \in [0, 1)$ is the capital share, $N \geq 0$ is the measure of workers that the firm employs, and $K \geq 0$ is the capital stock.

Firm-level productivity $z$ obeys a jump process. At Poisson rate $\chi > 0$, continuing firms draw a new productivity level from an exogenous distribution $\Gamma_0$. Firm productivity thus remains constant for a stochastic duration and then jumps to a new level that is independent from the previous one. I assume that $\Gamma_0$ satisfies the following conditions.

**Assumption 1:** The productivity distribution $\Gamma_0 : \mathbb{R}_+ \rightarrow [0, 1]$ is differentiable and satisfies

$$\Gamma_0'(z) > 0, \quad \int_0^\infty z \Gamma_0'(z) \, dz = 1, \quad \int_0^\infty z^{1-\alpha} \Gamma_0'(z) \, dz < \infty.$$

The first condition ensures that the density is strictly positive everywhere, the second condition is a normalization, and the third condition ensures that aggregate output is finite. For example, the third condition rules out Pareto distributions with exponents smaller than $1/(1-\alpha)$.

**Firm dynamics and matching.** At Poisson rate $\mu > 0$, an unemployed worker meets a potential firm. The potential firm then draws a productivity level $z \sim \Gamma_0$ and decides whether or not to enter. Continuing firms meet new workers at rate $\lambda N$, where the meeting rate $\lambda > 0$ is a technological parameter and $N$ is the measure of workers currently employed. This assumption represents a departure from the canonical Burdett and Mortensen (1998) model, who assume that firms meet workers at a rate that is independent of size. Burdett and Vishwanath (1988) call this linear meeting rate “balanced matching”. (In Section 5.3, I extend the model to allow for endogenous vacancy-posting and show that balanced matching arises as a limiting case.)

Matching is random, which implies that a fraction $u$ of meetings are with unemployed workers while the remainder $1-u$ are with employed workers, where $u$ is the unemployment rate. When a firm meets a worker, the worker only observes the wage posted by the firm and decides whether or not to accept the job offer. At Poisson rate $\delta \geq 0$, continuing workers exogenously separate to unemployment. At any moment, a firm can choose to exit and send all of its workers to unemployment.

Together, the assumptions of (i) constant return to scale in production, (ii) balanced matching, and (iii) the jump process for firm productivity allow me to obtain an analytical characterization of the stationary equilibrium of the model. As we will see shortly, the reason is that the firm problem becomes homogeneous (firm size is not a payoff relevant state variable) and that the stationary distribution is the solution to a quadratic equation.

**Strategies and beliefs.** When an employed worker receives an offer from a competing firm, she must form expectations about the future path of wages at both firms and choose the alternative with the highest continuation value. To make the analysis tractable, I restrict attention to Markov strategies that depend only on productivity. Denote by $w(z)$ the wage offered by a firm with productivity $z$, $e(z) \in \{0, 1\}$ the entry decision of a potential firm, and $x(z) \in \{0, 1\}$ the exit decision of a continuing firm. As in Coles and Mortensen (2016), I focus on equilibria where firm strategies do not depend directly on size. This is a natural choice given that the firm problem is homogeneous.

---

*A Poisson rate $\chi > 0$ means that the time elapsed between two events (in this case a productivity reset) is a random variable that is exponentially distributed with mean $1/\chi$.\footnote{A Poisson rate $\chi > 0$ means that the time elapsed between two events (in this case a productivity reset) is a random variable that is exponentially distributed with mean $1/\chi$.}
Distributions. Two related distributions will play a central role in my analysis: the employment-weighted productivity distribution \( P(z) \) and the wage distribution \( \tilde{P}(w) \). The employment-weighted productivity distribution summarizes the allocation of workers to firms of different productivity levels. Its definition is that, for a given productivity level \( z \), a share \( P(z) \in [0, 1] \) of workers are at firms with productivity less than \( z \). When more productive firms pay strictly higher wages (i.e., \( w'(z) > 0 \))—as will be the case in equilibrium—the following relationship holds:

\[
P(z) = \tilde{P}(w(z)).
\] (2)

In words, it says that the share of workers at firms with productivity less than \( z \) is equal to the share of workers earning a wage below \( w(z) \).

Worker problem. Denote by \( U \) the value of being unemployed and by \( W(w) \) the value of being employed at a firm currently paying wage \( w \). The value functions are the solutions to standard Bellman equations

\[
rU = b + \mu \int e(z) \left| W(w(z)) - U \right| d\Gamma_0(z) + \lambda (1 - u) \int \left| W(w(z)) - U \right| dP(z),
\]

where \( |x|_+ \equiv \max\{x, 0\} \), for any variable \( x \). The first term on the right-hand side of (3) accounts for the flow value of unemployment \( b \) while the second and third terms account for job offers by entering and continuing firms.

\[
rW(w) = w + \chi \int (1 - x(z))(W(w(z')) - W(w)) d\Gamma_0(z')
\]

wage changes due to productivity shocks

\[
+ \lambda (1 - u) \int \left| W(w(z')) - W(w) \right| dP(z') + \chi \int x(z) d\Gamma_0(z) + \delta \left( U - W(w) \right) dP(z).
\]

job offers by continuing firm  
job destruction

The first term on the right-hand side of (4) accounts for the flow of wages, the second term accounts for wage changes due to firm productivity shocks, the third term accounts for job offers by continuing firms, and the fourth term accounts for job destruction. Given that employed workers can quit to unemployment, the value of being employed must be weakly greater than the value of being unemployed:

\[
W(w) \geq U.
\] (5)

Proposition 1 characterizes the best response of workers to firms’ decision rules \( w(z) \), \( e(z) \), and \( x(z) \) along an equilibrium path.

**Proposition 1:** Unemployed workers accept a job offer if and only if it exceeds a reservation wage \( w \) given by

\[
w = b + \int \left( \mu e(z) - \chi (1 - x(z)) \right) \left( W(w(z)) - U \right) d\Gamma_0(z).
\] (6)

Employed workers accept a job offer if and only if it exceeds their current wage. They quit to unemployment if and only if their current wage falls below \( w \).
**Firm growth.** Equipped with the workers’ decision rule, I now characterize the link between firm-level wage and employment growth. The instantaneous change in employment at a firm of size $N$ paying $w \geq \bar{w}$ is given by

$$dN_t = \tilde{g}(w_t)N_t \, dt,$$

where the employment growth function of continuing firms $\tilde{g}(w)$ is an endogenous object that depends on the wage distribution $\tilde{P}(w)$. A simple expression can be obtained for the employment growth function.

**Lemma 1:** For all $w \geq \bar{w}$, the employment growth function is given by

$$\tilde{g}(w) \equiv \lambda u + \lambda (1 - u) \tilde{P}(w) - \lambda (1 - u)(1 - \tilde{P}(w)) - \delta.$$  \hspace{1cm} (8)

The function $\tilde{g}(\cdot)$ is weakly increasing and bounded from above by $\lambda - \delta$.

Lemma 1 highlights two important predictions of the model. First, high-wage firms grow faster. The reason is that offering a higher wage increases net poaching flows (i.e., the net amount of workers poached from other firms). Second, the marginal benefit of increasing the wage eventually converges to zero. To understand why, consider the case of a firm that offers the highest wage in the economy. Note that this firm is able to hire every worker that it meets and loses none of its workers to other firms. Hence, increasing the wage further would not lead to a higher growth rate of employment.

**Assumption 2:** The exogenous rates $(\lambda, \chi, \mu, \delta)$ satisfy

$$0 < \lambda - \delta < \min\{\chi, \mu\}$$

Assumption 2 is a sufficient condition for the existence of a steady-state unemployment rate and a well-defined employment-weighted productivity distribution (more on that in Section 2.2). The intuition is that if $\lambda - \delta \geq \min\{\chi, \mu\}$, some firms would grow too fast for too long. As a result, the stationary firm size distribution would have an infinite mean.

**Firm problem.** At every instant, the problem of a firm is to choose the current wage level $w$, the amount of capital per worker $k$, and whether to exit or continue operating. Given the constant return to scale assumption in both production and matching, the firm problem is homogeneous. Without loss of generality, I will characterize the value of a firm with a unit measure of employment $v(z) \equiv v(z, 1)$.

In the productivity region where the value of the firm is positive, the value function is the solution to the following Hamilton-Jacobi-Bellman (HJB) equation:

$$rv(z) = \max_{w \geq b, k \geq 0} \left\{ zk^\alpha - w - Rk + v(z)\tilde{g}(w) \right\} + \chi \left( \int v(x) \, d\Gamma_0(x) - v(z) \right).$$  \hspace{1cm} (9)

In the productivity region where the value of the firm is negative, the firm decides to exit.\footnote{More generally, the value function is the solution to a linear complementarity problem:

$$rv(z) \geq \max_{w \geq b, k \geq 0} \left\{ zk^\alpha - w - Rk + v(z)\tilde{g}(w) \right\} + \chi \left( \int v(x) \, d\Gamma_0(x) - v(z) \right).$$  \hspace{1cm} (10)}
The first-order conditions for capital and wages are respectively given by:

\[ z^{\alpha - 1} k^{\alpha - 1} = R, \]  \( (13) \)

\[ \frac{1}{m} = v'(z) \tilde{g}'(w). \]  \( (14) \)

The first-order condition for capital is standard and implies that firms rent capital up to the point where the marginal product of capital is equal to its user cost. The first-order condition for wages captures the trade-off between current profits and future growth. When a firm offers a high wage, it makes lower per-employee profit but grows faster, which increases the present value of future profits. The following proposition provides closed-form expressions for the policy functions.

**Proposition 2:** Capital per worker \( k(z) \), the wage schedule \( w(z) \), the entry decision \( e(z) \), and the exit decision \( x(z) \) are given by

\[ k(z) = \alpha^{\frac{1}{\alpha - 1}} R^{\frac{1}{\alpha - 1}} z^{\frac{1}{\alpha}}, \]  \( (15) \)

\[ w(z) = w + \int_{z}^{\infty} v(x) g'(x) \, dx, \]  \( (16) \)

\[ e(z) = 1 \{ z \geq z \}, \quad x(z) = 1 \{ z < z \} , \]  \( (17) \)

where \( z \geq 0 \) is a threshold. Capital per worker and wages are increasing in productivity \( z \).

Denote the equilibrium growth rate of employment by \( g(z) \equiv \tilde{g}(w(z)) \). Combining the expression for \( \tilde{g}(w) \) (see Equation 8) and the definition of the employment-weighted productivity distribution \( P(z) \) in the case where \( w'(z) > 0 \) (see Equation 2), I obtain

\[ g(z) = \lambda u + \lambda (1 - u) P(z) - \lambda (1 - u) (1 - P(z)) - \delta. \]  \( (18) \)

Labor productivity \( y(z) \) is defined as valued-added per worker while the labor share \( LS(z) \) is defined as payroll over value-added:

\[ y(z) \equiv zk(z)^\alpha , \quad LS(z) = w(z)/y(z). \]  \( (19) \)

### 2.2. Equilibrium

I now define and solve the equilibrium of the model. To simplify the notation, I define the flow of entering firms \( \mu_e \), the firm exit rate \( \chi_x \), the rate of arrival of productivity shocks \( \chi_s \),

\[ rv(z) \geq 0, \]  \( (11) \)

\[ 0 = v(z) \left( rv(z) - \max_{w \geq b, k \geq 0} \left\{ zk^\alpha - w - Rk + v(z) \tilde{g}(w) \right\} - \chi \left( \int v(x) d\Gamma_0(x) - v(z) \right) \right) . \]  \( (12) \)

The inequalities (10) and (11) ensure that the annuity value of a firm \( rv(z) \) is never less than the continuation value or the exit value. Equation 12 requires that either the value of a firm is zero, or that its annuity value is equal to its continuation value. Alternatively, the value function can be characterized as the solution to a variational inequality of the obstacle type (see Achdou, Buera, Lasry, Lions, and Moll, 2014).
and the productivity distribution of active firms \( \Gamma(z) \) as:

\[
\mu_e \equiv \mu(1 - \Gamma_0(\bar{z})), \quad \chi_x \equiv \chi \Gamma_0(\bar{z}), \quad \chi_s \equiv \chi(1 - \Gamma_0(\bar{z})), \quad \Gamma(z) \equiv \frac{\Gamma_0(z) - \Gamma_0(\bar{z})}{1 - \Gamma_0(\bar{z})}.
\]  

(20)

In Appendix A.4, I provide a derivation of the following laws of motion for the unemployment rate \( u \) and the employment-weighted productivity distribution \( P(z) \):

\[
\dot{u} = \left( \delta + \chi_x \right)(1 - u) - u(\mu_e + \lambda(1 - u)),
\]

(21)

\[
\dot{P}(z) = (1 - u)\lambda P(z)(P(z) - 1) + \frac{u}{1 - u} \mu_e \Gamma(z) + u\lambda P(z) - (\delta + \chi_x) P(z) + \chi_s (\Gamma(z) - P(z)).
\]

(22)

**Equilibrium definition.** A stationary equilibrium consists of a productivity distribution \( \Gamma(z) \), rates \( (\mu_e, \chi_x, \chi_s) \), value functions for the worker \((U, W(w))\), a reservation wage \( w \), a value function for the firm \( v(z) \), a productivity threshold \( \bar{z} \), decision rules for the firm \((w(z), k(z))\), an unemployment rate \( u \), and an employment-weighted productivity distribution \( P(z) \) that satisfy: (i) the truncated productivity distribution \( \Gamma(z) \) and rates \( \mu_e, \chi_x, \chi_s \) are determined by the optimal entry and exit of firms (20); (ii) the value functions for the worker \(U, W(w)\) solve (3, 4); (iii) the reservation wage \( w \) solves the worker problem (6); (iv) the value function \( v(z) \) and productivity threshold \( \bar{z} \) solve the linear complementarity problem (10, 11, 12); (v) the decision rules \( k(z) \) and \( w(z) \) solve the firm problem (15, 16); (vi) the unemployment rate \( u \) and the employment-weighted productivity distribution \( P(z) \) are the stationary solutions to their respective laws of motion (21, 22).

**Solution.** The following proposition provides an expression for the unemployment rate \( u \) and the employment-weighted productivity distribution \( P(z) \) in a stationary equilibrium.

**Proposition 3:** There exist a unique pair of unemployment rate \( u \) and employment-weighted distribution \( P(z) \) that are stationary solutions to their respective laws of motion (Equations 21 and 22):

\[
u = \frac{\lambda + \delta + \chi_x + \mu_e - \sqrt{\lambda + \delta + \chi_x + \mu_e}^2 - 4\lambda (\chi_x + \delta)}{2\lambda},
\]

(23)

\[
P(z) = \frac{\lambda(1-u) + \chi + \delta - u\lambda - \sqrt{\lambda(1-u) + \chi + \delta - u\lambda}^2 - 4\lambda(1-u)(\chi + \delta - u\lambda)\Gamma(z)}{2\lambda(1-u)}.
\]

(24)

The fact that wages are increasing in firm productivity (see Proposition 2) implies that the equilibrium exhibits rank preservation: the ranking of wages across firms is the same as the ranking of productivity. As in Burdett and Mortensen (1998) and Moscarini and Postel-Vinay (2013), it is the rank-preserving property of the model that yields tractability.

A corollary of Proposition 3 is that the allocation of workers to firm productivity ranks is invariant to the underlying productivity distribution \( \Gamma(\cdot) \). For example, the share of workers working at firms with productivity below \( z' \) depends only on \((\lambda, \chi, \delta, u)\) and the firm rank
\( \Gamma(z') \in [0,1] \), but not on the absolute productivity level \( z' \). The feature of the equilibrium that generates this result is that the optimal quit decision of worker is ordinal: they accept any job offer that generates a wage increase.

**ASSUMPTION 3:** *The value of unemployment \( b \) is high enough so that \( v(z) = 0 \)*

From now on, I will focus on the case where the value of unemployment \( b \) is “high enough”. For low values of \( b \), it could be that firms with \( z = 0 \) can profitably operate (i.e., \( v(0) > 0 \)), which would imply that there is no firm exit. Assumption 3 contrasts with Coles and Mortensen (2016), who assume that the lower bound of the productivity distribution is such that all firms make positive profits in every date and state, which implies \( v(z) > 0 \).

2.3. *Micro and macro predictions*

I now derive predictions regarding the link between (i) productivity and labor share at the firm level and (ii) productivity dispersion and the aggregate labor share. Before doing so, I provide intuition regarding the equilibrium relationship between wages and productivity at the firm level.

**Intuition.** The first-order condition for capital (13) implies that the share of payments to capital in value-added \( Rk(z)/y(z) \) is equal to the output elasticity of capital \( \alpha \). The equilibrium wage \( w(z) \) thus determines how the remaining \( 1 - \alpha \) share of value-added is split between wages and profits. The first-order condition for wages (14) evaluated at the equilibrium wage yields the following equation for the pass-through of productivity to wages:

\[
w'(z) = g'(z) \times v(z).
\]

When a firm’s productivity increases, the marginal increase in the wage offered is exactly equal to the marginal increase in the value of the firm. The value of the firm increases because (i) higher wages imply a higher growth rate (the growth effect) and (ii) the firm extracts rents from every additional worker (the value effect). Using Equation 18, the growth effect can be further decomposed into two terms

\[
g'(z) = 2\lambda(1-u) \times P'(z).
\]

First, search frictions mitigate the effect of wages on net poaching flows. For instance, the growth effect is lower if the meeting rate \( \lambda \) is low or the unemployment rate \( u \) is high. Second, the extent to which a wage increase yields higher net poaching flows depends on the density of workers working at firms with productivity \( z \). At the margin, a wage increase only affects the quit behavior of workers who were initially indifferent between staying or quitting. Since \( P'(z) \) is a density, it must be that \( P'(z) \rightarrow 0 \), so firms in the upper tail of the productivity distribution face vanishing wage competition and therefore have no incentive to pass productivity gains to their workers.

The fact that high-productivity firms tend to make higher profits has been discussed in Bon temps et al. (2000) in the context of the standard Burdett-Mortensen model. A unique prediction of my model— which arises due to the interaction between productivity shocks and the endogenous exit decision—is that low-productivity firms pay wages that are *above* the static marginal
product of labor. Re-arranging the Bellman equation (9), the value to the firm of a worker can be decomposed into the flow profits and an option value:

\[
v(z) = \frac{1}{r + \chi_{\pi} - g(z)} \left[ (1 - \alpha)y(z) - w(z) + \chi_s \left( \int v(x) \, d\Gamma(x) - v(z) \right) \right]. \tag{27}
\]

Evaluating at \( z = \bar{z} \) and using the fact that \( v(\bar{z}) = 0 \) (Assumption 3), I obtain an expression for the equilibrium wage of the lowest productivity firm:

\[
w = (1 - \alpha)y(z) + \chi_s \int_0^{\infty} v(x) \, d\Gamma(x). \tag{28}
\]

Notice that workers at firms with \( z = \bar{z} \) extracts all the surplus from the match, including the option value of future productivity increases. As a result, they earn a wage that exceeds their static marginal product.

**Firm-level labor shares.** The following proposition characterizes the relationship between productivity and labor share at the firm level.

**Proposition 4—Micro:** There exists a threshold \( z' \) such that firms with \( z < z' \) have a labor share above the frictionless benchmark (i.e., \( \text{LS}(z) > 1 - \alpha \)). There exists a second threshold \( z'' \) such that for firms with \( z > z'' \), the labor share is decreasing in productivity (i.e., \( \text{LS}'(z) < 0 \)).

Figure 1a presents the relationship between the labor share and labor productivity in a numerical example. In contrast with the frictionless model where the labor share would be \( 1 - \alpha \) for every firm, the model predicts that the labor share of high-productivity firms is lower than \( 1 - \alpha \), while the opposite is true for low-productivity firms.

\[
\text{(a) Micro} \quad \text{(b) Macro}
\]

**Figure 1.—Relationship between labor share and productivity (numerical example)**
**Aggregation.** The aggregate labor share is equal to the output-weighted average of firm-level labor shares. I now use the analytical solution of the stationary equilibrium to obtain an expression for the aggregate labor share.

**PROPOSITION 5—Aggregation:** The aggregate labor share is given by

\[
\text{LS} = \int_{\tilde{z}}^{\infty} \frac{\chi - \tilde{g}}{\chi - g(z)} \times \frac{y(z)}{Y} \times \frac{\text{employment share}}{\text{relative productivity}} \times d\Gamma(z),
\]

where \(\tilde{g} \equiv \int_{z}^{\infty} g(z) \, dP(z)\) is the average growth rate of continuing firms and \(Y \equiv \int_{z}^{\infty} y(z) \, dP(z)\) is aggregate output.

Notice that output shares are increasing in productivity \(z\). The reason is twofold. First, high-productivity firms tend to employ more workers (i.e., their employment share is higher). Second, they produce more output per worker (i.e., their relative productivity is higher). The aggregate labor share is thus disproportionately determined by the labor share of high-productivity firms.

To study the link between the distribution of productivity \(\Gamma(z)\) and the aggregate labor share, I now restrict attention to a tractable special case that yields a closed-form solution. In particular, I assume that the discount rate \(r\) is equal to zero and that the distribution of productivity is Pareto (i.e., \(\Gamma_0(z) = 1 - z^{-\frac{1}{\sigma}}\)). The parameter \(\sigma\) governs the thickness of the right tail of the productivity distribution. Since the variance of the logarithm of productivity \(z\) is precisely \(\sigma\), I refer to \(\sigma\) as “productivity dispersion”. Assumption 1 requires that \(\sigma < \frac{1}{1 - \alpha}\).

**PROPOSITION 6—Pareto special case with no discounting:** Suppose that \(\Gamma_0(z) = 1 - z^{-\frac{1}{\sigma}}\) for \(z \geq 1\) and \(r = 0\). The aggregate labor share is given by

\[
\text{LS} = 1 - \alpha - \frac{\chi - \tilde{g}}{\chi - g(z)} \times \frac{\text{profit share}}{\text{output share}},
\]

where \(\frac{\chi - \tilde{g}}{\chi - g(z)} \in (0, 1)\), which implies that \(\text{LS} \in (0, 1 - \alpha)\).

The key takeaway is that, all else constant, the aggregate labor share is decreasing in productivity dispersion \(\sigma\). The Cobb-Douglas exponent \(\alpha\) determines the share of payments to capital while productivity dispersion \(\sigma\) is a key determinant of how the remaining \(1 - \alpha\) share of aggregate value-added is split between wages and profits. Notice that free-entry exhausts ex-ante rents, so that without productivity dispersion (i.e., \(\sigma = 0\)), there would be no ex-post profits.\(^9\)

---

\(^8\)In a slight abuse of language, I refer to the object \(\omega(z) \equiv \int_{z}^{\infty} \frac{\chi - \tilde{g}}{\chi - g(z)} \times \frac{y(z)}{Y} \times \frac{\text{employment share}}{\text{output share}} \times d\Gamma(z)\) as the output share of firms with productivity \(z\). To be precise, \(\omega(z)\) is a density that satisfies \(\int \omega(z) \, d\Gamma(z) = 1\). Similarly, I refer to \(\int_{z}^{\infty} \frac{\chi - \tilde{g}}{\chi - g(z)} \times \frac{y(z)}{Y} \times \frac{\text{employment share}}{\text{output share}} \times d\Gamma(z)\) as the employment share of firms with productivity \(z\).

\(^9\)To be precise, the marginal entrant makes zero profits in expectation (i.e., \(v(z) = 0\) as per Assumption 3). Therefore, if all firms had the same productivity level (i.e., \(\sigma = 0\)), then the present value of aggregate profits would be zero. In a stationary equilibrium with no discounting, this implies zero aggregate profits.
To understand why the aggregate labor share is decreasing in productivity dispersion, recall from (26) that the pass-through of productivity to wages is proportional to the amount of local competition that a firm faces (i.e., the density of workers at firms with a similar productivity level). An increase in productivity dispersion $\sigma$ makes firms become “further apart” in terms of productivity, effectively weakening between-firm competition in the labor market.

Finally, notice that the persistence of the productivity process $\chi$ affects both the level of the aggregate labor share as well as its sensitivity to productivity dispersion $\sigma$. All else equal, if productivity differentials between firms were purely transitory (i.e., $\chi \to +\infty$), the labor share would converge to the frictionless case $\text{LS} = 1 - \alpha$, where aggregate profits are zero. In that case, productivity dispersion would not affect the labor share. If instead productivity differentials across firms were permanent (i.e., $\chi = 0$), then the labor share would converge to $1 - \alpha - \sigma$. Therefore, what matters for the share of profits in GDP is not only how far apart firms are in terms of productivity, but how persistent these productivity differentials are.

**Comparison with benchmark models.** I now contrast the prediction of the baseline model with two benchmark models. First, I consider an extension of the model in (Burdett and Mortensen, 1998) with capital (henceforth “Burdett-Mortensen”). The main difference with the baseline model is that productivity differentials are permanent instead of being mean-reverting. Second, I consider a variation of the Coles and Mortensen (2016) model with capital and balanced matching (henceforth “Coles-Mortensen”). The only difference with the baseline model is that entry and exit is exogenous. Note that the original model developed in Coles and Mortensen (2016) does not have balanced matching, but instead has a convex hiring cost function. I study the implications of costly hiring (i.e., endogenous vacancy-posting) separately in Section 5.3.

**Table I**

<table>
<thead>
<tr>
<th>Model</th>
<th>Aggregate labor share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frictionless</td>
<td>$1 - \alpha$</td>
</tr>
<tr>
<td>Baseline model</td>
<td>$1 - \alpha - \frac{\chi x - \overline{g}}{\chi} \sigma$</td>
</tr>
<tr>
<td>Coles-Mortensen</td>
<td>$1 - \alpha - \left(\frac{\chi x - \overline{g}}{\chi} + \frac{\chi x - \overline{g}}{\chi - \overline{g}} \right) \sigma$</td>
</tr>
<tr>
<td>Burdett-Mortensen</td>
<td>$1 - \alpha - \sigma$</td>
</tr>
</tbody>
</table>

**Notes:** The models are ranked according to their aggregate labor share, from highest to lowest. See Appendix B for a formal derivation of the Coles-Mortensen and Burdett-Mortensen models. “Productivity dispersion” $\sigma$ is the inverse of the Pareto exponent; $\overline{g} \equiv \lambda u - \delta$ is the average growth rate of continuing firms; $\overline{g} \equiv -\lambda (1 - 2u) - \delta$ is the growth rate of firms with $z = \tilde{z}$.

Table I contains closed-form expressions for the aggregate labor share in the Pareto special case with no discounting for each benchmark model. First, notice that the labor share in the baseline model is higher than in the Coles-Mortensen model. The reason is that, in the baseline model, endogenous entry exhausts rents up to the point where the marginal entrant has a continuation value of zero (i.e., $v(\tilde{z}) = 0$). In Coles-Mortensen, entry is exogenous and all firms make weakly positive flow profits, which implies that $v(\tilde{z}) > 0$. As a result, profits are higher in Coles-Mortensen and the labor share is lower.

Second, the labor share in Burdett-Mortensen is the lowest of all models considered. To understand why, notice that the labor share in the Burdett-Mortensen model coincides with the labor share in the baseline model when there is no mean-reversion in productivity (i.e., $\chi = 0$).
As discussed earlier, a high persistence of productivity differentials reduces the labor share in favor of profits.

The main takeaway from Table I is that both endogenous entry and mean-reversion in firm productivity reduce the amount of rents that accrue to firms. As a result, these model ingredients generate a higher labor share (i.e., closer to the frictionless benchmark) and a lower sensitivity of the labor share to productivity dispersion. Figure 1b contains a numerical example where the aggregate labor share is shown as a function of productivity dispersion (keeping everything else constant).

While the results in Table I are obtained under strong assumptions (i.e., a Pareto distribution and zero discounting), they provide theoretical guidance as to which model parameters are important: the capital share $\alpha$, the productivity process $(\sigma, \chi)$, and the pace of job reallocation between firms $(\bar{g}, \chi_x)$. In the next section, I conduct a quantitative analysis of the model where all functional forms and parameter values are disciplined by moments in the data.

3. QUANTIFYING THE THEORY

3.1. Calibration

I calibrate the model to the Canadian economy over the 2000–2015 period. As in the US, the Canadian labor share is now lower than it was in the 1980s but, unlike in the US, the Canadian labor share appears to have been quite stable since 2000 (see Appendix D.2). I will thus compare the 2000–2015 period to the stationary equilibrium of the model.

Data. I use microdata from the National Accounts Longitudinal Microdata File (henceforth “NALMF”), which is constructed by merging administrative data from different sources. The NALMF contains de-identified data covering the universe of private sector employers in Canada over the 2000–2015 period. The unit of observation is an enterprise, which is an entity larger than an establishment but smaller than a consolidated corporation. Unlike with establishments, these entities have a full set of income statement and balance sheets, which makes it possible to construct firm-level measures of value-added.

I restrict the main sample to the private corporate sector excluding: Agriculture, Mining, Utilities, Education, and Health (i.e., NAICS 11, 21, 22, 61, and 62). Agriculture and Mining are excluded due to data limitations while Utilities, Education, and Health are excluded due to the fact that these sectors are dominated by public entities in Canada. Finally, I restrict the sample to firm-year observations with more than 5 employees that have no missing values in employment, value-added, capital stock, and industry code.

The data allows me to compute value-added at the firm level using the same methodology as in the System of National Accounts. Using the income approach, I construct value-added as the sum of worker compensation (i.e., income that accrues to workers) and gross profit (i.e., income that accrues to firm owners). I can then compute the labor share as the ratio of worker compensation to value-added. Labor productivity is defined as the ratio of value-added to employment. In Appendix D.1, I describe the variable construction in more details and provide summary statistics. The final sample contains 3,084,182 firm-year observations, with an average of 192,761 firms per year.

10The full definition taken from Statistics Canada’s website is “Enterprise refers to the highest level of the Business Register statistical hierarchy and is associated with a complete set of financial statements. The enterprise, as a statistical unit, is defined as the organizational unit of a business that directs and controls the allocation of resources relating to its domestic operations, and for which consolidated financial and balance sheet accounts are maintained from which international transactions, an international investment position and a consolidated financial position for the unit can be derived. It corresponds to the institutional unit as defined for the System of National Accounts.”
**Productivity distribution.** I model the productivity distribution $\Gamma_0(z)$ as a Gamma distribution with a mean normalized to one. The Gamma distribution has both a thick upper tail and a cusp, which is consistent with the typical shape of productivity distributions in the data. The resulting distribution is fully characterized by a single shape parameter $\eta > 0$ that governs the thickness of the upper tail, where a lower $\eta$ implies a thicker tail.\footnote{The PDF is defined over $\mathbb{R}_+$ and is given by $\Gamma_0(z) = \frac{\eta^\eta}{G(\eta)} z^{\eta-1} e^{-\eta z}$, where $G(\eta)$ is the Gamma function.} As I show in Appendix C.3, the Gamma distribution provides a good fit for the empirical distribution of productivity, while the Pareto distribution assumed in Proposition 6 generates too much skewness. In Section 5.4, I consider an alternative calibration of the baseline model where $\Gamma_0(z)$ is a Pareto distribution.

**Externally calibrated parameters.** The model has nine parameters. I externally calibrate the first three $(r, d, \alpha)$ and internally calibrate the last six $(\lambda, \mu, \delta, b, \chi, \eta)$. I interpret a time period $[t, t + 1)$ in the model to be a year. The value of $r$ is chosen to imply an annual discount rate of 5%. For the depreciation rate $d$, I obtain a target by using industry-level data from Statistics Canada (Table 031-0006, linear depreciation). For the set of industries covered in the main sample, the average annual depreciation rate over the 2000–2015 period is 14.8%. Finally, I set $\alpha = 0.230$ to match the capital-output ratio of 1.10 measured in the NALMF data. Table II summarizes these choices. To map discrete-time growth rates $g_d$ in the data to continuous-time growth rates $g_c$ in the model, I use the conversion formula $g_c = \log(1 + g_d)$.\footnote{Suppose that a continuous-time process satisfies $\dot{x}_t = g_c x_t$. Solving forward, we have that $\frac{x_{t+1}}{x_t} = e^{g_c}$. The definition of discrete-time growth rate is $g_d \equiv x_{t+1}/x_t - 1$. Combining both equations and solving for $g_c$, I obtain $g_c = \log(1 + g_d)$.}

**TABLE II**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest rate</td>
<td>$r$</td>
<td>0.049</td>
<td>Annual discount rate of 5%</td>
</tr>
<tr>
<td>Depreciation rate</td>
<td>$d$</td>
<td>0.160</td>
<td>Annual depreciation rate of 14.8%</td>
</tr>
<tr>
<td>Capital share</td>
<td>$\alpha$</td>
<td>0.230</td>
<td>Capital stock to annual output ratio of 1.10</td>
</tr>
</tbody>
</table>

**Internally calibrated parameters.** The remaining parameters $\theta \equiv (\lambda, \mu, \delta, b, \chi, \eta)'$ are jointly calibrated. I target moments from the data and choose the set of parameters that minimizes the distance between the model-implied moments and the data. (In Appendix C.3, I derive analytical formulas for the model-implied moments.) To ensure that Assumption 2 is satisfied, I restrict the parameters space to $\Theta \equiv \mathbb{R}_+^6 \cap \{\theta \mid 0 < \lambda - \delta < \min\{\chi, \mu\}\}$. I target 7 moments related to the dynamics of firm-level productivity and the pace of job reallocation between firms. The moments are: the job creation rate by continuers, the job creation rate by entrants, the job destruction rate by continuers, the job destruction rate by exiters, the unemployment rate, the autocorrelation of firm-level log labor productivity, and the interdecile range of log labor productivity.

I take the unemployment rate, job creation rate by entrants, and job destruction rate by exiters directly from publicly available data (Statistics Canada tables 282-0087 and 527-0001) and compute 2000–2015 averages. The job creation rate by entrant is defined as the sum of all jobs created by entering firms in a year divided by the total number jobs in the economy at
the beginning of the year. The job destruction rate by exiters is defined as the number of jobs destroyed by exiting firms divided by the number of total jobs in the economy.

The reason why I do not use the NALMF microdata to estimate the job creation rate (destruction rate) by entrants (exiters) is because my main sample is not exactly the universe of firms due to sample restrictions. Therefore, firms moving in and out of the sample does not correspond to firm entry and exit. Moreover, it is not possible to accurately distinguish entry and exits from mergers and acquisitions using only the NALMF data.

The remaining moments are computed using the NALMF microdata. The job creation rate by continuers is defined as the sum of jobs created by expanding firms divided by the number of total jobs in the economy. Similarly, the job destruction rate by continuers is defined as the sum of jobs destroyed by shrinking firms divided by the number of total jobs in the economy. The autocorrelation of firm-level log labor productivity is obtained by estimating the autoregressive parameter in a firm-level panel AR(1) regression with year and 3-digit NAICS fixed effects. To compute the interdecile range of log labor productivity, I first remove year and (3-digit NAICS) industry fixed effects from firm-level log labor productivity. I then compute the difference between the 90th and 10th percentiles of the (residualized) log labor productivity in the pooled 2000–2015 sample.

**Numerical solution.** In Appendix C.1, I provide a detailed description of the solution method. Denote the vector of moments in the data by $\hat{\Lambda}$. For each value of $\theta \in \Theta$, I solve the model numerically and compute the model-implied moments $\Lambda(\theta)$. I then choose the parameter $\theta$ that minimizes the following quadratic form

$$
\min_{\theta \in \Theta} \left( \Lambda(\theta) - \hat{\Lambda} \right)' \left( \Lambda(\theta) - \hat{\Lambda} \right).
$$

(31)

Since $\Theta$ is not compact, I implement the numerical optimization problem by using its closure $\tilde{\Theta}$ and then verify that the minimizer $\theta^*$ is in $\Theta$. I obtain the parameter estimate $(\lambda, \mu, \delta, b, \chi, \eta) = (0.2443, 0.3063, 0.0233, 0.5981, 0.2311, 3.8773)$. Table III contains the fit of the model and in Appendix C.3, I report the sensitivity of the targeted moments to changes in parameter values (see paragraph titled “Jacobian matrix”).

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment rate</td>
<td>0.071</td>
<td>0.071</td>
</tr>
<tr>
<td>Autocorrelation of log labor prod.</td>
<td>0.810</td>
<td>0.806</td>
</tr>
<tr>
<td>Interdecile range of log labor prod.</td>
<td>1.547</td>
<td>1.547</td>
</tr>
<tr>
<td>Job creation (continuers)</td>
<td>0.061</td>
<td>0.055</td>
</tr>
<tr>
<td>Job creation (entrants)</td>
<td>0.019</td>
<td>0.022</td>
</tr>
<tr>
<td>Job destruction (continuers)</td>
<td>0.067</td>
<td>0.062</td>
</tr>
<tr>
<td>Job destruction (exiters)</td>
<td>0.016</td>
<td>0.016</td>
</tr>
</tbody>
</table>

The only structural parameters that can be interpreted directly are the flow value of unemployment $b$ and the difference $\lambda - \delta$. The average wage in the economy is 1.0606, so $b$ is equal to 56% of the average wage, which is close to (but higher than) the value of 0.4 typically used in the labor search literature. Recall from Lemma 1 that the annual growth rate of employment of a firm is bounded from above by $e^{\lambda - \delta} - 1 \approx 0.25$. The calibration therefore implies that
firms can not grow faster than 25% annually. Strictly speaking, this is counterfactual given that the empirical firm growth distribution typically contains very high values (Bottazzi and Secchi, 2006). Nevertheless, the constraint imposed by search frictions in the model is not overly restrictive. For example, a firm that remains at the top of the productivity distribution for 25 years will grow its employment by a factor of $e^{25 \times (\lambda - \delta)} \approx 250$.

3.2. Validation

I now validate the model by assessing its ability to match the data along several dimensions. Given that the calibration strategy does not use data on wages or the labor share, everything from now on is non-targeted.

**Aggregate and average labor share.** As a starting point, consider the empirical counterpart of Equation 29

$$LS = \sum \omega_i LS_i,$$

which states that the aggregate labor share is the sum of firm-level labor shares $LS_i$, weighted by their value-added share $\omega_i$. Re-arranging, the aggregate labor share can be expressed as the sum of the average (unweighted) firm-level labor share and the covariance between labor share and value-added share

$$LS = \bar{LS} + \text{cov}(LS_i, \omega_i).$$

In Table IV, I report the three components of Equation 33 in the data and in the model. In the data, I compute the components separately in each 2-digit NAICS industry and then take a value-added weighted average. I then compute a simple average over the 2000–2015 period.

<table>
<thead>
<tr>
<th>Model</th>
<th>Aggregate</th>
<th>Average</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>0.653</td>
<td>0.884</td>
<td>-0.231</td>
</tr>
<tr>
<td>Baseline model</td>
<td>0.616</td>
<td>0.889</td>
<td>-0.273</td>
</tr>
<tr>
<td>Coles-Mortensen</td>
<td>0.559</td>
<td>0.547</td>
<td>+0.011</td>
</tr>
<tr>
<td>Burdett-Mortensen</td>
<td>0.533</td>
<td>0.512</td>
<td>+0.021</td>
</tr>
</tbody>
</table>

The model generates an aggregate labor share of 0.62, which is close but slightly lower than in the data (0.65). The average (unweighted) firm-level labor share in the model is 0.89, which is also close to the data (0.88). The resulting covariance between labor share and value-added share is negative, and stands at $-0.27$ in the model ($-0.23$ in the data). Overall, the model replicates the main features of the data: a high average labor share and a negative covariance between value-added and labor share in the cross-section of firms.

Table IV also contains the same decomposition for the Coles-Mortensen and Burdett-Mortensen models (see Appendix B for a description of the calibration strategy). Consistent with the theoretical result in the Pareto special case (see Table I), the labor share is lower in
the Coles-Mortensen and Burdett-Mortensen economies. Notice that, in both cases, the covariance between value-added share and labor share is close to zero, while it is strongly negative in the data and in the baseline model. As we will see shortly, this is due to the fact that the benchmark models can not generate labor shares above $1 - \alpha$ (i.e., negative flow profits) for low-productivity firms.

Assessing cross-sectional moments. The negative cross-sectional covariance between labor share and value-added share is not only a feature of the Canadian data. It has previously been documented in the case of the Taiwanese manufacturing sector (Edmond et al., 2015), the U.S. manufacturing sector (Kehrig and Vincent, 2021), and most other sectors of the U.S. economy (Autor et al., 2020). The value-added share of a firm is high either because it is large (it employs a lot of workers) or because it it has a high labor productivity (it produces a lot of value-added per worker). A key prediction of the model is that the negative relationship between value-added and labor share is entirely driven by productivity, not size (see Proposition 2). I now confront this prediction with the data.

Table V contains results from a panel regression of firm-level labor shares on different covariates (all variables are in logarithm) with year and 3-digit NAICS industry fixed effects. Specification (1) regresses labor share on labor productivity. I obtain a negative coefficient of $-0.31$. As a point of comparison, I compute the (population) regression coefficient in the calibrated model and obtain a value of $-0.62$. One concern with the regression specification is that the presence of measurement error in value-added could lead to a spurious negative relationship between labor share and labor productivity. In Appendix E.1, I address this concern by instrumenting labor productivity with its one-year lag. I obtain a slightly lower value in absolute terms ($-0.25$).

Specification (2) regresses firm-level labor shares on both labor productivity and size. The coefficient on labor productivity remains mostly unchanged ($-0.31$) while the coefficient on employment is close to zero ($< 0.01$). This is roughly in line with the model predictions, where (i) introducing employment as a control does not affect the coefficient on labor productivity and (ii) the coefficient on employment is zero. Specification (3) regresses firm-level labor shares on size by itself. The relationship is negative but very small ($-0.001$). In the model, the coefficient is larger in absolute value but also very small ($-0.007$). Through the lens model, the relationship between employment and labor share is spurious in the sense that it only reflects the positive correlation between productivity and employment in a stationary equilibrium.

The empirical relationship between labor productivity and labor share could in principle be explained by the fact that firms differ in their capital share (i.e., the share of payments to capital in value-added). For example, if a firm purchases a robot to replace some of its workers, then labor productivity will increase (because the number of employees decreases) and the share of payments to capital in value-added will increase (because the stock of capital increases). In Appendix E.1, I estimate Specification (1) controlling for the capital-output ratio and find a negligible role for (within-industry) capital intensity heterogeneity in explaining differences in firm labor shares across firms.

Overall, I find robust support for the idea that high-productivity firms have a lower labor share. In the model, productivity and wages have a rank correlation of one (i.e., high-wage firms are necessarily high-productivity firms). Hence, one would expect that the relationship between average wage (i.e., compensation per worker) and labor share be positive. When I estimate

---

13 The labor share is defined as the ratio of worker compensation to value-added while labor productivity is defined as value-added per worker. If there is a measurement error in value-added, there will be a spurious negative correlation between value-added and labor share.
### TABLE V
ASSESSING CROSS-SECTIONAL MOMENTS

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Model</td>
<td>Data</td>
</tr>
<tr>
<td>Labor share (log LS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor productivity (log y)</td>
<td>−0.313*** (0.002)</td>
<td>−0.620 (0.002)</td>
<td>−0.313***</td>
</tr>
<tr>
<td>Employment (log N)</td>
<td>0.009*** (0.001)</td>
<td>0</td>
<td>−0.001 (0.001)</td>
</tr>
<tr>
<td>Industry fixed effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sample size</td>
<td>3,084,182</td>
<td>3,084,182</td>
<td>3,084,182</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.272</td>
<td>0.273</td>
<td>0.036</td>
</tr>
</tbody>
</table>

Notes: Standard errors in parentheses are clustered at the firm-level (*** $p < 0.01$).

this cross-sectional relationship in the data, I instead obtain a weak positive relationship. In Appendix E.2, I use an accounting decomposition to understand why the model fails to match this particular moment. In short, the reason is that the model does not generate enough wage dispersion between firms.

**Unpacking the aggregate labor share.** I now assess whether the model-implied relationship between labor productivity, value-added share, and labor share is quantitatively consistent with the microdata. Denote by $LS_d$ and $\omega_d$ the labor share and value-added share within a labor productivity decile $d$. The aggregate labor share can be expressed as

$$LS = \sum \omega_d LS_d.$$  \hspace{1cm} (34)

I construct the labor productivity deciles in the data by sorting firm-year observations by labor productivity within industry-year bins. I then pool industries and years together and compute the measure of interest within each decile (see Appendix D.3 for details).

The left panel of Figure 2 contains the labor share by labor productivity deciles in the model and in the data. First, notice that there are large differences in the labor share across firms. In the model, the difference between the labor share at the top decile versus the bottom decile is 1.22 (1.14 in the data). To provide a benchmark, I include a horizontal line at $1 - \alpha$, which corresponds to the labor share that would prevail in a frictionless competitive model. Notice that a large fraction of firms have a high labor share. In the model and in the data, the bottom 5 deciles of firms have a labor share above $1 - \alpha$. In contrast, the top decile of firms have a low labor share of 0.52 (0.44 in the data). In line with the predictions of the baseline model (see Proposition 4), labor shares above $1 - \alpha$ are a pervasive feature of the data. The Burdett-Mortensen and Coles-Mortensen models instead predict that all firms have a labor share strictly below $1 - \alpha$ (see Appendix Figure E.1d).

Both in the model and in the data, value-added is extremely concentrated within high-productivity firms (see right panel of Figure 2). In the model, the top decile of firms account for 55% of of value-added (42% in the data), while the bottom 5 deciles accounts for a mere 11% (18% in the data). Through the lens of the model, the negative covariance between labor share and value-added share documented earlier can thus be explained by the fact that labor
shares are decreasing in firm productivity while value-added shares are (on average) increasing in firm productivity.

In Appendix E.3, I conduct a number of robustness checks to assess the importance of industries, firm size, capital share heterogeneity, and measurement error in value-added. Overall, I find that the main findings are robust: (i) there is a strong negative relationship between firm-level labor share and labor productivity and (ii) a large fraction of firms have a labor share above \(1 - \alpha\).

Assessing panel moments. I now assess the ability of the model to match panel moments (i.e., the comovement of variables within firms over time). I focus on two core predictions of the model: (i) firms respond to an increase in productivity by increasing their wage (see Proposition 2), which (ii) induces an increase in their employment growth (see Lemma 1).

To compare model and data, I use the following auxiliary models:

\[
\log w_{i,t+k} = \mu_i^k + \lambda_t^k + \beta_w^k \log y_{i,t} + u_{i,t+k},
\]

\[
\Delta^k \log N_{i,t} = \mu_i^k + \lambda_t^k + \beta_g^k \log w_{i,t} + u_{i,t+k}.
\]

The parameters \(\mu_i\) and \(\lambda_t\) denote firm and year fixed effects. The variables \(y_{i,t}\), \(w_{i,t}\), and \(\Delta^k \log N_{i,t}\) denote, respectively, labor productivity, average wage, and \(k\)-year ahead log employment growth for firm \(i\) in year \(t\). The coefficient \(\beta_w^k\) represents the response of wages at time \(t + k\) to an increase of labor productivity at time \(t\). Similarly, \(\beta_g^k\) is the response of employment growth between \(t\) and \(t + k\) to a wage increases at time \(t\).

Figure 3a plots the dynamic response of wages to an increase in labor productivity over a five-year horizon, both in the model and in the data.\(^{14}\) In the data, I do not observe worker-level wages, so I construct a measure of the average wage of a firm by dividing total worker compensation by employment. On impact, wages increase less than one-for-one with labor productivity (\(\beta_w^0 = 0.40\) in the model and \(\beta_w^0 = 0.57\) in the data). The interpretation is that, on average, workers capture roughly half of productivity increases through higher wages. After four years, wages have fully reverted back to zero both in the model and in the data, reflecting the fact that productivity shocks are mean-reverting.

\(^{14}\)Using the calibrated model, I simulate synthetic trajectories of employment, wages, and labor productivity for 200,000 firms. I first simulate 16 years of monthly data, and then aggregate to the annual frequency. I obtain a panel of synthetic data, which I use to estimate (35) and (36) exactly as in the NALMF data.
(a) Labor productivity shock

(b) Wage shock

Figure 3.—Assessing panel moments (coefficients from specifications 35 and 36)

Figure 3b plots the dynamic response of cumulative employment growth to an increase in wages. Both in the model and in the data, cumulative employment growth increases strongly over the first two years and then stabilizes. At the one-year horizon, the interpretation is that when a firm increases its wage by 10%, its annual growth rate of employment increases by 3.7 percentage points in the model and 1.9 in the data. The model thus overpredicts the employment response to a wage increase. Taken at face value, this divergence between model and data means that the model does not generate enough monopsony power. To see that, notice that the first-order condition for wages (14) implies that $1 = v(z)\tilde{g}'(w)$. Therefore, a high response of employment growth to a wage change $\tilde{g}'(w)$ must coexist with a low present value of rents accruing to firms (i.e., a low level of monopsony power).

The relationship between labor productivity and wages that I estimate is related to existing empirical evidence on the wage-productivity pass-through. The empirical literature consistently finds pass-through elasticities that are positive and well below one (see Card, Cardoso, Heining, and Kline, 2018 for a literature review). Similarly, my estimates of the response of employment growth to wages is related to the literature studying separation elasticities (see Manning, 2011 for a literature review).  

It is worth noting that my dynamic estimates correspond to correlations in the data and do not necessarily have a causal interpretation. To address this, I confront the model with existing causal evidence. Kline, Petkova, Williams, and Zidar (2019) is the most relevant study since it estimates both the elasticity of labor productivity to wages (pass-through elasticity) and the elasticity of worker separations to wages (separation elasticity) in the cross-section of U.S. firms. The identification strategy is based on comparing firms whose patent applications were initially allowed to those whose applications were initially rejected. They find that firms respond to rising productivity by increasing worker compensation, which induces a decline in worker separations. In Table VI, I replicate the estimates from Kline et al. (2019) in the model. The model matches both moments closely: the pass-through elasticity in the model is slightly higher ($0.54$ versus $0.47$) and the separation elasticity is slightly lower in absolute value ($-1.49$ versus $-1.62$).

The existing literature mostly studies the effect of wages on the separation rate rather than employment growth (which is the difference between the hiring rate and the separation rate). One potential reason is that the hiring rate is often thought to depend not only on the wage, but also on recruiting intensity (usually modelled as vacancy-posting in search model). In Section 5, I study an extension of the model with endogenous vacancy-posting.
TABLE VI  
MODEL PREDICTIONS AND EVIDENCE FROM KLINE ET AL. (2019)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass-through elasticity</td>
<td>0.47</td>
<td>0.54</td>
</tr>
<tr>
<td>Separation elasticity</td>
<td>$-1.62$</td>
<td>$-1.49$</td>
</tr>
</tbody>
</table>

Notes: The pass-through elasticity is obtained from Column 2 of Table VIII. It is defined as the percentage change in wage bill per worker in response to a 1% change in value-added per worker. The separation elasticity is obtained from Column 1 of Table IX. It is defined as the percentage change in the fraction of workers who separate during the next year in response to a 1% change in the wage. In Appendix C.3, I provide expressions for the analogous model-implied moments.

**Wage distribution.** One of the key predictions of wage-posting models is that wage dispersion can be sustained in equilibrium (i.e., the law of one price does not hold). As a final validation exercise, I compare the wage distribution in the model and in the data. I use data from Bowlus, Gouin-Bonenfant, Liu, Lochner, and Park (2021) on the distribution of annual earnings in Canada over the 2000–2015 period and compare it with the distribution of wages across workers in the calibrated model.

The variance of log wages is 0.73 in the data, while it is only 0.15 in the model (see Table VII). Most noticeably, the model can not explain the amount of right-tail inequality in the data: the P99-P90 percentile difference is 1.32 in the data, while it is only 0.34 in the model. Since the model does not include worker heterogeneity, this is not surprising. Following the work of Abowd, Kramarz, and Margolis (1999)—who use two-way fixed effect regressions (i.e., AKM regression) to decompose earnings inequality across workers—researchers have found that “worker fixed effects” are the key determinant of earnings inequality.

TABLE VII  
WAGE DISTRIBUTION.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log wage distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total variance</td>
<td>0.73</td>
<td>0.15</td>
</tr>
<tr>
<td>P99-P90</td>
<td>1.32</td>
<td>0.34</td>
</tr>
<tr>
<td>P90-P50</td>
<td>0.78</td>
<td>0.71</td>
</tr>
<tr>
<td>P50-P10</td>
<td>0.75</td>
<td>0.29</td>
</tr>
<tr>
<td>Variance of firm fixed effects</td>
<td>0.07</td>
<td>0.15</td>
</tr>
<tr>
<td>Firm-size wage premium</td>
<td>0.03</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Notes: Data on the wage distribution is obtained from Bowlus et al. (2021) (annual earnings, 2000–2015 average); Firm-size wage premium is obtain by regression firm-level average wage on employment (both in logarithms, $N = 3,084,182$).

Both in the US and in Canada, recent studies employing an AKM framework have found that firms account for roughly 10% of the variance of log earnings (see Song, Price, Guvenen, Bloom, and Von Wachter, 2019 for the U.S. and Dostie, Li, Card, and Parent, 2020 for Canada). According to this measure, the model fits the data much more closely: the variance of log wages attributable to firms is 0.15 in the model versus 0.07 in the data.  

16In Section 4, I consider an extension of the model with endogenous recruiting intensity that improves the fit of the model. In particular, the model-implied variance of log wages goes down from 0.15 to 0.12.
firm-size wage premium in the model and in the data. As in the data, the model predicts a small premium: firms with 10% more employees pay wages that are on average 0.1% higher (0.3% higher in the data).

4. PRODUCTIVITY DISPERSION AND THE LABOR SHARE

I now use the calibrated model to simulate the effect of a rise in productivity dispersion.

Motivating evidence. To motivate the exercise, I first present data on productivity dispersion and the labor share in Canada and the U.S. over the 2000–2015 period (see Table VIII). For comparability, I use publicly available data on the labor share from Statistics Canada and the BEA. In both countries, I compute productivity dispersion as in Section 3.1, but using sales per worker as a proxy for labor productivity.\(^{17}\) I use Compustat data for the U.S. and the NALMF for Canada. In Appendix D.4, I validate my estimate of productivity dispersion for the U.S. against Census microdata from Barth, Bryson, Davis, and Freeman (2016).

<table>
<thead>
<tr>
<th>Year</th>
<th>Labor share</th>
<th>Productivity dispersion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Canada</td>
<td>United States</td>
</tr>
<tr>
<td>2001–2005</td>
<td>0.646</td>
<td>0.675</td>
</tr>
<tr>
<td>2006–2010</td>
<td>0.646</td>
<td>0.650</td>
</tr>
<tr>
<td>2011–2015</td>
<td>0.650</td>
<td>0.616</td>
</tr>
</tbody>
</table>

Change (%) 0.4 −5.8 0.01 0.23

Notes: The labor share is calculated using BEA table 1.14 (corporate sector) in the U.S. and Statistics Canada Table 3800063 in Canada. In both cases, I compute the aggregate labor share as the ratio of “compensation of employees” to the sum of “compensation of employees” and “gross operating surplus”. Productivity dispersion is calculated using Compustat data for the U.S. and the NALMF for Canada. In both cases, I compute the interdecile range of (log) sales per worker net of 2-digit NAICS and year fixed effects, excluding industries 11, 21, 61, and 62.

One striking observation is that over the 2000–2015 period, the labor share has declined by nearly 6 percentage point in the U.S. while it has remained essentially unchanged in Canada. The mirror image is true for productivity dispersion: the interdecile range of sales per worker has increased by 23 log points in the U.S. but by only 1 log point in Canada.

4.1. Model experiment

I now use the model to simulate the effect of a rise in productivity dispersion of the same magnitude as what the U.S. economy as experienced over the 2000–2015 period. To do so, I lower the parameter \( \eta \) from 3.8773 to 2.9974, which implies a rise in the interdecile range of labor productivity increases by 23 log points. All other parameters remain unchanged.

The model experiment consists of comparing the stationary equilibrium of the model before and after the shock. Table IX summarizes the effect of this “productivity dispersion shock”. First, the aggregate labor share declines by 1.3 percentage points. To put into perspective, a

\(^{17}\)The reason why I use sales per worker instead of labor productivity is that firm-level value-added data does not exist in the U.S. outside of the manufacturing sector.
permanent 1.3% decline in the labor share is quite large, representing roughly a quarter of the 5.8% decline of the U.S. labor share reported in Table VIII. In my framework, the capital-output ratio would need to rise by 6.3% in order to generate such a decline of the labor share.\textsuperscript{18} Interestingly, the average (unweighted) labor share moves in the opposite direction, increasing by 7.5%. The capital share remains unchanged, so the decline of the labor share is exactly compensated by a rise of the profit share. The rising wedge between average and aggregate labor share implies that the covariance between labor share and value-added share increases (in absolute value).

Recall that the productivity distribution $\Gamma_0(z)$ has a mean normalized to one, so a change in its shape parameter $\eta$ represents a mean-preserving spread. However, given that employment is concentrated within high-productivity firms, the mean-preserving spread generates an increase in aggregate output of 7.3%. The unemployment rate remains essentially unchanged, reflecting the fact that the rise in productivity dispersion has little effect on firm entry and exit. (This is consistent with the Jacobian matrix in Appendix Table C.II, which indicates that the only targeted moment that is meaningfully responsive to a change in $\eta$ is productivity dispersion.)

*Adjustment and reallocation effects.* To highlight the forces at play, I now study the impact of the productivity dispersion shock along the productivity distribution. Figure 4 plots the response of labor shares and value-added shares by labor productivity deciles. The positive response of the average (unweighted) labor share reported in Table IX can be explained by the fact that, for all but the top 3 deciles, firm labor shares increase. In contrast, the labor share of high-productivity firms decrease.

The intuition is that firms in the right tail of the productivity distribution become further apart in terms of productivity, which effectively shields them from wage competition. As a result, they keep a larger share of value-added as profits. On the other end of the spectrum, low productivity firms are now willing to absorb larger losses in order to retain their workers. The reason is that the rising profit margins of high-productivity firms lead to an increase of the option value of retaining workers.

While employment shares remain unchanged (see Proposition 3), the relative productivity of firms at the top of the distribution increases, so their value-added share increases. Figure 4b indicates that the bottom 9 deciles of firms lose value-added shares at the expense of the top decile. Taking stock, the model predicts that most firms see an upward adjustment of their labor

\begin{table}[h]
\centering
\caption{Response of aggregates (model experiment)}
\begin{tabular}{lcccr}
\hline
Variable & Steady-state & Counterfactual & Change (pp.) \\
\hline
Labor share & 0.616 & 0.603 & -1.33 \\
Average labor share & 0.889 & 0.964 & 7.50 \\
Covariance & -0.273 & -0.361 & -8.83 \\
Capital share & 0.230 & 0.230 & 0.00 \\
Profit share & 0.154 & 0.167 & 1.33 \\
Output & 1.598 & 1.714 & 7.26 \\
Unemployment & 0.071 & 0.071 & -0.01 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{18}In my calibration, the user cost of capital is $R = 20.9\%$ and the capital share $R \times \frac{K}{Y}$. If $\frac{K}{Y}$ increases by 6.4\%, the labor share would decline by $0.209 \times 6.4\% \approx 1.3\%$.
share, but that there is a reallocation of value added towards high-productivity firms who have a low labor share.

To quantify the relative importance of the adjustment and reallocation components, I use a shift-share decomposition as in Baily, Hulten, and Campbell (1992). Using Equation 34, the change in the aggregate labor share $\Delta LS$ can be expressed as the sum of three terms:

$$\Delta LS = \sum \omega_d \Delta LS_d + \sum \Delta \omega_d LS_d + \sum \Delta \omega_d \Delta LS_d,$$

where $\omega_d$ and $LS_d$ denote the value-added share and labor share of firms in productivity decile $d$. The adjustment term measures the effect of changing the labor shares $LS_d$ while keeping the value-added shares $\omega_d$ constant. It accounts for a small positive increase in the aggregate labor share ($+0.03\%$). The reallocation term measures the effect of changing the value-added shares $\omega_d$ while keeping the labor shares $LS_d$ constant. The reallocation component is the dominant force pushing the labor share down ($-1.08\%$). Finally, Kehrig and Vincent (2021) suggest a “directed reallocation” interpretation of the third term. It accounts for the comovement of labor shares and value-added shares. Since there is reallocation of value-added towards firms with a declining labor share, the directed reallocation term is negative ($-0.29\%$).

**Related empirical evidence.** The model provides a rich set of predictions regarding the effect of a productivity dispersion shock and many of them can be related to existing empirical studies. First, the decline of the labor share in the model is driven by a reallocation of value-added towards firms with a low (and declining) labor share, and is partially offset by an increase in the labor share of most firms. This is precisely what Kehrig and Vincent (2021) document in the U.S. manufacturing sector. They show that over the 1967–2012 period, the aggregate labor share declined by 4.5 percentage point per decade, while the median labor share increased by 0.7 percentage point per decade. They reconcile this divergence by showing that low-labor-share firms have gained value-added shares over time. Over this same period, Kehrig (2015) documents a secular increase in TFP dispersion in the U.S. manufacturing sector. Similarly,
Autor et al. (2020) study all U.S. industries and find that falling labor shares are largely accounted for by reallocation rather than a fall in the (unweighted) average labor share across firms.

Second, the model predicts that, in response to a productivity dispersion shock, the concentration of value-added increases while employment concentration remains constant. Consistent with this prediction, Kehrig and Vincent (2021) find that there has been a massive reallocation of value-added shares towards low-labor-share firms, but that there was limited reallocation of employment. At the aggregate level, there has been a secular increase in sales concentration in the U.S. over the past two decades (Grullon et al., 2019) that has coincided with the decline of the labor share. The negative comovement between labor share and concentration at the industry-level is often interpreted as evidence in favor of rising monopoly power (Barkai, 2020, Autor et al., 2020). In my model, however, a rise in productivity dispersion generates a negative comovement of these two variables despite the fact that concentration has no causal effect on the labor share.

4.2. Supportive evidence

The existing literature highlights the fact that there has been substantial heterogeneity in labor share dynamics across countries (see Gutiérrez and Piton, 2020) and industries (see Autor et al., 2020). How much of this can be explained by differential trends in productivity dispersion? To answer this question, I now estimate the relationship between labor share and productivity dispersion using cross-industry and cross-country evidence. Specifically, I estimate the following specifications

\[
\text{LS}_{j,t} = \mu_t + \beta \sigma_{j,t} + \epsilon_{j,t},
\]

(38)

\[
\text{LS}_{j,t} = \mu_t + \gamma_j + \beta \sigma_{j,t} + \epsilon_{j,t},
\]

(39)

\[
\Delta^5 \text{LS}_{j,t} = \mu_t + \gamma_j + \beta \Delta^5 \sigma_{j,t} + \epsilon_{j,t+5},
\]

(40)

where \(\sigma\) denotes productivity dispersion (i.e., the interdecile range of log labor productivity), \(\text{LS}\) denotes the labor share, \(\mu_t\) is a year fixed effect, \(\gamma_j\) is a unit fixed effect (industry or country depending on the dataset used), and \(\Delta^5\) denotes a five-year difference.

The first specification (Equation 38) uses cross-sectional variation of levels for identification. The second specification (Equation 39) uses variation in \(\text{LS}\) and \(\sigma\) within an industry/country over time for identification. For the last specification (Equation 40), I use non-overlapping 5-year periods to calculate changes, so the identification comes from the low frequency comovement of \(\text{LS}\) and \(\sigma\). Barkai (2020) and Autor et al. (2020) both use a version of specification (40) to estimate the effect of concentration on the labor share.

Cross-industry evidence. The relative stability of productivity dispersion at the aggregate level in Canada hides a lot of heterogeneity at the industry level. I now leverage this cross-industry heterogeneity to estimate the relationship between labor share and productivity dispersion. To do so, I aggregate the NALMF microdata to construct a panel dataset covering most 3-digit NAICS industries over the 2000–2015 period (see Appendix D.5 for summary statistics).

If the Canadian labor market was perfectly segmented along industry lines (i.e., firms within an industry only compete with each other for workers), then each industry could be treated as a separate labor market, and the model predictions would be directly comparable to the cross-industry evidence. In practice, however, workers often move across industries over their career
(see Jarosch et al., 2019 for empirical evidence from Austria). Nevertheless, one would expect the labor share in an industry to be negatively related to its productivity dispersion as long as the labor market is at least partially segmented along industry lines. (In Appendix B.4, I make this point formally in a simplified version of the baseline model.)

In line with the model experiment, I find that industries with high (rising) productivity dispersion systematically have a low (declining) labor share. Panels (1), (2), and (3) of Table X regress industry-level labor shares on productivity dispersion using specifications (38), (39), and (40). In all specification, I obtain negative coefficients ranging from $-0.063$ (specification 39) to $-0.118$ (specification 39). As a frame of reference, the implied slope in the model experiment is $-0.0133/0.23 = -0.058$. Figures 6a and 6b contain scatter plots summarizing the relationship between labor share and productivity dispersion (in levels and in 10-year changes).

### TABLE X
**CROSS-INDUSTRY REGRESSIONS**

<table>
<thead>
<tr>
<th>LS</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>$-0.112^{***}$</td>
<td>$-0.118^{**}$</td>
<td>$-0.063^*$</td>
<td>$-0.094^{***}$</td>
<td>$-0.096^{***}$</td>
<td>$-0.071^*$</td>
</tr>
<tr>
<td></td>
<td>$(0.029)$</td>
<td>$(0.047)$</td>
<td>$(0.034)$</td>
<td>$(0.026)$</td>
<td>$(0.023)$</td>
<td>$(0.020)$</td>
</tr>
<tr>
<td>HHI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-0.137$</td>
<td>$-0.533^{***}$</td>
<td>$-0.041^{***}$</td>
<td>$(0.218)$</td>
<td>$(0.071)$</td>
<td>$(0.095)$</td>
</tr>
<tr>
<td>K/Y</td>
<td>$-0.029$</td>
<td>$-0.022$</td>
<td>$-0.026$</td>
<td>$(0.218)$</td>
<td>$(0.071)$</td>
<td>$(0.095)$</td>
</tr>
<tr>
<td>Year FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Industry FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$N$</td>
<td>1104</td>
<td>1104</td>
<td>207</td>
<td>1104</td>
<td>1104</td>
<td>207</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.154</td>
<td>0.736</td>
<td>0.228</td>
<td>0.188</td>
<td>0.776</td>
<td>0.331</td>
</tr>
</tbody>
</table>

*Notes: Standard errors in (1), (2), (4), and (5) are clustered at the industry level (**p < 0.01, **p < 0.05, *p < 0.1). “LS” denotes labor share; “$\sigma$” denotes productivity dispersion (i.e., the interdecile range of log labor productivity); “K/Y” denotes the capital to output ratio. For specifications (3) and (6), the windows are 2000–2005, 2005–2010, and 2010–2015.*

Other factors such as capital deepening (Karabarbounis and Neiman, 2014) and concentration (Barkai, 2020; Autor et al., 2020) have been show to affect the labor share. To control for these other factors, I reproduce specifications (38), (39), and (40) by adding the capital-output ratio and the HHI (sum of squared firm-level sales shares) as controls (see panels (4), (5), and (6) of Table X). Consistent with existing findings, the coefficients on the HHI and the capital-output ratio are all negative. However, the addition of controls does not materially affect my main findings. If anything, the effect of productivity dispersion becomes more precisely estimated—i.e., lower standard errors and less variation across specifications—and the estimated slopes are closer to the model-implied value of $-0.058$.

I now provide a more stringent evaluation of the model’s mechanism. In the model experiment, a rise in productivity dispersion leads to a decline of the aggregate labor share that operates through a reallocation of value-added towards high-productivity firms, as opposed to a broad-based decline of firm-level labor shares. In each industry and year, I sort firms into labor productivity quintiles. For each productivity quintile $q$, I estimate the following regression

$$Y_{q,j,t} = \mu_t + \gamma_j + \beta_q \sigma_{j,t} + \varepsilon_{q,j,t}, \quad (41)$$
where $\mu_t$ is a year fixed effect and $\gamma_j$ is a 3-digit NAICS industry fixed effect. $Y_{q,j,t}$ represents the variable of interest (labor share or value-added share) in productivity quintile $q$, industry $j$, and year $t$.

Figure 5 plots the estimated coefficients with their 95% confidence intervals (standard errors clustered at the industry level). In panel (a) we can see that the labor share of low-productivity firms increases with productivity dispersion while the opposite is true for high-productivity firms. The coefficients are imprecisely estimated, but for the two top quintiles they are negative and significant at the 5% level. To compare with the model predictions, Figure 5 contains model-implied slopes, which are computed by dividing the change of the variable of interest (i.e., labor share of value-added share) within a specific productivity quintile by the change in productivity dispersion in the model experiment.

Turning to the response of value-added shares (panel (b) of Figure 5), the coefficient for the top quintile of firms is positive while it is negative or near-zero for all the other quintiles. Only the top quintile is significant at the 5% level. Taken as a whole, these findings are remarkably consistent with the transmission mechanism in the calibrated model.

Cross-country evidence. Finally, I repeat the exercise using cross-country data. For the cross-country exercise, I use harmonized data on productivity dispersion produced by the OECD.\(^{19}\) The measure of productivity dispersion in the dataset is exactly the same as I used for the cross-industry exercise (i.e., the interdecile range of log labor productivity in the cross-section of firms). The measures are computed at the NACE two-digit industry level and then averaged with employment weights to provide, for each country-year, a measure for the manufacturing sector and for non-financial services sectors. I merge the productivity dispersion data with publicly-available data on sectoral labor shares.\(^{20}\) I obtain a balanced panel of 7 countries.

\(^{19}\)The OECD has developed a project called MultiProd that seeks to provide harmonized cross-country data on productivity and wage dispersion. The data they collect is obtained by running a standardized routine on individual country’s production surveys and business registers. The variables that I use (productivity dispersion) is taken directly from the replication material of Berlingieri, Blanchenay, and Criscuolo (2017).

\(^{20}\)The data on the labor share by sector is obtained from EUKLEMS project (Denmark, Finland, France and Italy), WORLDKLEMS (Japan), and directly from national statistics institutes (Norway and New Zealand). For a description
Figure 6.—Productivity dispersion and the labor share across countries and industries

Notes. In all panels, the x-axis is “Productivity dispersion”, defined as the the interdecile range of log labor productivity in the cross-section of firms and the y-axis is “Labor share”. In each figure, the OLS slope is reported in the top right corner. Panels (a) and (b) present the data in level for the year 2001. Panels (b) and (c) contain 10-year differences of both variables over the 2001–2011 period. For panels (a) and (b), 3-digit NAICS industries are sorted into broad sectors: “Goods-producing” includes industries part of NAICS codes 23 and 31–33; “Trade” includes industries part of NAICS codes 41,48, and 49; “Professional services” includes industries part of NAICS codes 51–55; “Other services” includes industries part of NAICS codes 44, 45, 56, 71, 72, and 81.

(Denmark, Finland, France, Italy, Japan, Norway and New Zealand) and two broad industry groups (manufacturing and non-financial services) for the 2001–2011 period. One year of data was missing for France and Japan, so I used linear interpolation.

Figures 6c and 6d contain scatter plots summarizing the relationship between labor share and productivity dispersion (in levels and in 10-year changes). Consistent with the model predictions, I find that countries with high (rising) productivity dispersion systematically have a low (declining) labor share. In Appendix E.4, I also report the regression results using specifications (38), (39), and (40).
5. EXTENSIONS

I now consider a number of model extensions in order to assess how sensitive the results of the model experiment in the previous section are to different modelling choices. Overall, I find that the baseline model is somewhat conservative in the sense that most of the model extensions considered imply a larger labor share decline in response to a rise in productivity dispersion. I organize the results in Table XI, which contains the steady-state aggregate labor share across model extensions as well as the changes in output, unemployment rate, and labor share induced by a rise in productivity dispersion (i.e., a 23 log point increase in the interdecile range of labor productivity, exactly as in Section 4).

<table>
<thead>
<tr>
<th>Model</th>
<th>Steady-state</th>
<th>Experiment (% change)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LS</td>
<td>Output</td>
</tr>
<tr>
<td>Benchmark models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline model</td>
<td>0.616</td>
<td>+7.26</td>
</tr>
<tr>
<td>Coles-Mortensen</td>
<td>0.559</td>
<td>+7.30</td>
</tr>
<tr>
<td>Burdett-Mortensen</td>
<td>0.533</td>
<td>+10.2</td>
</tr>
<tr>
<td>Model extensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline + endogenous b</td>
<td>0.616</td>
<td>+7.30</td>
</tr>
<tr>
<td>Baseline + endogenous λ</td>
<td>0.622</td>
<td>+15.0</td>
</tr>
<tr>
<td>Baseline + endogenous (λ, b)</td>
<td>0.622</td>
<td>+14.7</td>
</tr>
<tr>
<td>Pareto distribution</td>
<td>0.387</td>
<td>+24.7</td>
</tr>
<tr>
<td>Pareto distribution + endogenous b</td>
<td>0.387</td>
<td>+25.2</td>
</tr>
</tbody>
</table>

5.1. Benchmark models

First, I replicate the model experiment using the calibrated Burdett-Mortensen and Coles-Mortensen models described in Appendix B. Consistent with the theoretical results in the Pareto special case with no discounting (see Table I), the change in the labor share induced by an increase in productivity dispersion in the baseline model (−1.3%) is smaller in absolute value than in the Coles-Mortensen model (−1.6%), which itself is smaller than in the Burdett-Mortensen model (−2.0%).

5.2. Endogenous-b extension

In the baseline model, there is small decline of the unemployment rate in response to a rise in productivity dispersion (see Table IX). The reason is that the equilibrium entry threshold $z$ responds to the increased profitability associated with becoming a highly-productive firm. In contrast, the benchmark models do not have that force (i.e., firm entry is exogenous).

For complete comparability with the benchmark models, I now consider an extension of the baseline model where I shut down the equilibrium response of the unemployment rate (henceforth “endogenous-b extension”). To do so, I allow the flow value of unemployment $b$ to depend on all the model parameters $\theta$. In particular, I assume that $b(\theta)$ is such that the
equilibrium unemployment rate is always 0.0713 (i.e., the unemployment rate in the steady-state of the baseline model). 21

Repeating the model experiment, I find that the response of the labor share in the endogenous-\(b\) extension is essentially the same as in the baseline model (see Table XI). The reason is that the unemployment rate response in the baseline model is negligible (−0.01 percentage point). As we will see shortly, shutting down the response of the unemployment rate will be more important in other model extensions.

5.3. Endogenous-\(\lambda\) extension

The balanced matching assumption in the baseline model implies that the rate at which firms meet workers is exogenous. As a result, the only way that firms can increase their hiring rate is by increasing their wage. I now consider a model extension where firms can increase the rate at which they meet workers by posting more vacancies (henceforth “endogenous-\(\lambda\) extensions”).

Environment. From now on, I interpret the meeting rate \(\lambda\) in the baseline model as a vacancy rate (i.e., the measure of vacancies per employee). I assume that firms choose \(\lambda\) optimally subject to a convex cost function \(c(\lambda, z)\). The employment growth function (Equation 18 in the baseline model), is now given by

\[
g(\lambda, w) = \frac{\lambda}{\tilde{f}(w)} \times \tilde{f}(w) - (\Lambda(1 - \tilde{Q}(w)) + \delta),
\]

where the “vacancy yield” is equal to the probability that a job offer is accepted (i.e., \(\tilde{f}(w) \equiv u + (1 - u)\tilde{P}(w)\)). As in the baseline model, \(\tilde{P}(w)\) denotes the wage distribution across workers. The new object \(\Lambda\) denotes the aggregate vacancy rate and \(\tilde{Q}(w)\) denotes the vacancy-weighted wage distribution. The key difference with the baseline model is that firms can increase their hiring rate either through a higher vacancy rate \(\lambda\), or through a higher wage, which increases the vacancy yield \(\tilde{f}(w)\).

For reasons that will become clear shortly, I consider an isoelastic vacancy-posting cost function of the form

\[
c(\lambda, z) = \frac{\lambda^{-\theta} v(z)}{1 + \theta} \lambda^{1+\theta}.
\]

The parameters \(\lambda, \theta > 0\) govern, respectively, the level and convexity of the vacancy-posting cost function. The scaling factor \(\frac{1}{1 - \alpha}\) ensures that the capital share is equal to \(1 - \alpha\) for all firms, as in the baseline model. Notice that the cost function is proportional to the present value of a hire \(v(z)\). The interpretation is that firms with a vacancy rate \(\lambda\) have to pay a flow of real resources equal to a fraction \(\frac{\lambda^{-\theta} v(z)}{1 + \theta} \lambda^{1+\theta}\) of their value as “vacancy-posting costs”.

---

21To be precise, the flow value of unemployment \(b\) is now a function of all the model parameters \(\theta\) and the equilibrium unemployment rate is a function of exogenous parameters \(\theta\) as well as \(b\). The function \(b(\theta)\) is thus implicitly defined as \(u(\theta, b(\theta)) = 0.0713\).
**Equilibrium.** As in the baseline model, wages $w(z)$ are increasing in firm productivity and firm entry and exit is determined by an endogenous threshold $z$. The following proposition highlights the equilibrium behavior of firms as it relates to hiring.

**Proposition 7:** The equilibrium vacancy yield and vacancy rate are given by

$$f(z) = u + (1 - u)P(z),$$  \hfill (Vacancy yield)

$$
\lambda(z) = \frac{\partial f(z)}{\partial z}. \hfill (Vacancy rate)
$$

The equilibrium vacancy yield $f(z) \equiv \tilde{f}(w(z))$ is increasing in firm productivity $z$, owing to the fact that wages are increasing in productivity. The vacancy rate $\lambda(z)$ is also increasing in productivity. The reason is that a high vacancy yield induces high-productivity firms to post more vacancies per worker. Notice that the parameter $\lambda$ governs the level of vacancy-posting while the parameter $\theta$ determines the sensitivity of the vacancy rate to firm productivity. The baseline model is nested as a special case where deviation from $\lambda(z) = \lambda$ are infinitely costly (i.e., $\theta \to \infty$), which implies that all firms have the same vacancy rate.

**Calibration.** To calibrate the new parameter $\theta$, I use external evidence on firm vacancies from Davis, Faberman, and Haltiwanger (2013) (henceforth DFH). Using establishment-level data on vacancies in the US over the 2002–2006 period, DFH shows that high-growth firms have both a higher vacancy rate (i.e., number of vacancies per employee) and a higher vacancy yield (i.e., number of hires per vacancy). However, they estimate that the vacancy yield is much more elastic, with respect to the hiring rate, than the vacancy rate. The following proposition clarifies how I map the DFH evidence to the model.

**Proposition 8:** The equilibrium cross-sectional elasticities of the vacancy rate $\lambda$ and vacancy yield $f$ to the hiring rate $h$ are given by:

$$\frac{d \log \lambda}{d \log h} = \frac{1}{1 + \theta}, \quad \frac{d \log f}{d \log h} = \frac{\theta}{1 + \theta}.$$

Consider two extreme cases. If the vacancy cost function is linear (i.e., $\theta = 0$), then differences in hiring rates across firms are entirely due to differences in the vacancy rate (i.e., $d \log \lambda / d \log h = 1$). If instead, the cost function is infinitely convex (i.e., $\theta \to +\infty$), then differences in hiring rates across firms are entirely due to differences in the vacancy yield (i.e., $d \log f / d \log h = 1$). In short, $\theta$ determines the relative importance of the vacancy rate and the vacancy yield in explaining the heterogeneity in hiring rates across firms.

DFH estimates a robust log-linear relationship between the vacancy yield and the hiring rate with an elasticity of $d \log f / d \log h = 0.82$ (see Figure IX in Davis et al., 2013). The choice of the cost function (43) is thus justified ex-post by the fact that it generates an exact log-linear relationship between these two variables. The idea of backward engineering the cost function to match the evidence from DFH has previously been used in Gavazza, Mongey, and Violante (2018).

For the calibration, I first set $\theta = 0.82 / 1 = 0.82$ and then jointly calibrate the model parameters $(\lambda, \theta, \mu, \delta, b, \chi, \eta)$ by targeting the same 7 moments as in the baseline model. In Appendix C.1, I extend the solution algorithm to account for endogenous vacancy-posting and in Appendix C.4, I report the targeted moment fit.
Quantitative results. The aggregate labor share in the endogenous-λ extension is about one percentage point higher than in the baseline model (see Table XI), entirely due to a lower profit share (the capital share is \( \alpha = 0.23 \) in both models, see Appendix B.3 for a discussion). Moreover, the aggregate labor share responds somewhat more mildly to a productivity dispersion shock than in the baseline model (see Table XI). Shutting down the equilibrium response of the unemployment rate (i.e., row “endogenous-(\( \lambda, b \)) extension” of Table XI), the response is slightly lower in absolute value.

These finding suggests that the addition of a vacancy-posting margin effectively reduces the amount of monopsony power. A related finding emerges in Manning (2006), who studies a simple model where monopsonistic firms can raise employment either by increasing the wage offered or increasing expenditures on recruiting. The author shows that the convexity of the recruiting cost function is a central determinant of the amount of monopsony profits.

Quantitatively, I find that moving from \( \theta \rightarrow \infty \) (in the baseline model) to \( \theta = 4.56 \) (in the endogenous-λ extension), implies only a modest decrease in the profit share of about 1 percentage point. To understand why, it is important to go back to the micro moment that pins down \( \theta \). Figure 7a plots the equilibrium hiring rate against the vacancy yield, both in logarithms. While the endogenous-λ extension has a slope lower than in the baseline model (0.82 versus 1), the slope is still well above 0, indicating that high-growth firms use the wage margin intensively to increase their hiring. Through the lens of the model, it must mean that the vacancy-posting cost function has a lot of curvature (i.e., \( \theta \) is high).

Finally, It is worth pointing out that the addition of a vacancy-posting margin has the additional effect of reducing wage dispersion. Figure 7b plots the (relative) wage distribution in the baseline model and in the endogenous-λ extension. The distribution becomes visibly less spread out, and the variance of log wages across workers declines from 0.15 to 0.12, putting it more in line with the empirical target of 0.07 (see “variance of firm effects” in Table VII).

5.4. Pareto distribution

As a final robustness check, I now consider an alternative calibration strategy where the productivity distribution \( \Gamma_0(z) \) is Pareto as assumed in Proposition 6. To ensure that the calibration strategy is fully consistent with the baseline model, I parametrize the Pareto distribution so that
its mean is normalized to one. In particular, I assume that

$$\Gamma(z) = 1 - (1 - \sigma)^{-\frac{1}{\delta}} z^{-\frac{1}{\delta}} \quad \text{for} \quad z \geq 1 - \sigma,$$

where the tail index satisfies $\sigma < \frac{1}{1 - \alpha}$ to ensure finite aggregate output (see Assumption 1).

I solve the model and calibrate it using the same algorithm as for the baseline model. In Appendix C.4, I report the targeted moment fit. It is worth noting that the Pareto model fails to match the (non-targeted) joint distribution of labor productivity and labor share. In particular, it generates a labor share for firms in the top decile of labor productivity that is much lower than in the data (0.23 versus 0.53, see Appendix Figure E.1e). Relatedly, it implies a low aggregate labor share (0.39 versus 0.62 in the baseline model, see Table XI).

Repeating the model experiment, I find that the response to the productivity dispersion shock is much stronger than in the baseline model (i.e., $-2.2\%$ versus $-1.3\%$, see Table Table XI). Shutting down the equilibrium response of the unemployment rate (i.e., row “Pareto distribution + endogenous-$b$ extension” of Table XI), the response is even larger, at $-2.4\%$.

6. CONCLUDING REMARKS

In this paper, I study the effect of labor market imperfections on the labor share in a tractable firm dynamics model with search frictions. In the model, heterogeneous firms grow by accumulating workers and compete through wages in a frictional labor market. In equilibrium, high-productivity firms pay wages below the marginal product while low-productivity firms pay wages above the marginal product. I calibrate the model using administrative data covering the universe of firms in Canada over the 2000–2015 period and show that it can replicate a number of non-targeted features of the data, despite being very parsimonious.

The key theoretical finding is that the distribution of firm productivity is a central determinant of the aggregate labor share. I find that productivity dispersion effectively weakens the intensity of wage competition in the labor market. Quantitatively, the model predicts that a rise in productivity dispersion of the same magnitude as what the U.S. economy has experienced over the 2000–2015 period causes a 1.3 percentage point decline in the aggregate labor share. Importantly, the decline of the labor share operates via a reallocation of value-added towards firms with a low (and declining) labor share, thus providing a new interpretation to recent firm-level evidence on the decline of the U.S. labor share (see Autor et al., 2020; Kehrig and Vincent, 2021). Exploiting heterogeneity in labor share trends across countries and industries, I provide regression evidence in support of the idea that countries and industries with high (rising) productivity dispersion systematically have a low (declining) labor share.

The analysis in this paper has focused on firm heterogeneity while abstracting from worker heterogeneity. Future work should attempt to quantify which workers benefit the most from working at high-productivity firms. For instance, are “superstar workers” able to extract their full marginal product when working at “superstar firms”?

REFERENCES


