Skill Requirements, Search Frictions and Wage Inequality*

Lawrence Uren  
University of Melbourne  

Gabor Virag  
University of Rochester  

June 21, 2009

Abstract

We analyze wage inequality in the context of a Burdett-Mortensen (1998) model extended to incorporate worker heterogeneity through skill requirements in the production process. Equilibrium wage dispersion is a consequence of worker and firm production heterogeneity and frictions in the search process. We provide sufficient conditions for more productive firms to offer higher wages and characterize this equilibrium when it exists. The model is calibrated to match observed unemployment and the distribution of firm productivity and we find the introduction of skill requirements improves the ability of the model to match the data. Finally, we use the model to explain the joint movement of both within- and between-group inequality in the late 1980s and 1990s, movements which other models have difficulty explaining.

*The authors would like to thank Arpad Abraham, Roland Benabou, Jeff Borland, Andrew Clarke, Chris Edmond, John Kennan, Ian King, Thomas Lemieux, Ashley Lester, Ian McDonald, Jonathan Parker, Wolfgang Pesendorfer, and Robert Shimer. A working paper version of this paper was circulated under the title “Wage Inequality in a Burdett-Mortensen World.” We are also grateful to participants of seminars at the Australian National University, the University of Melbourne, and the Summer Economics Conference at the Hungarian Academy of Science for useful comments and discussions. All remaining errors are ours.
1 Introduction

Wage inequality has increased dramatically in the United States over the last few decades.\(^1\) This change is reflected in movements in both between- and within-group inequality. Throughout most of the 1980s, between- and within-group inequality increased simultaneously. However, this pattern changed in the late 1980s. While between-group inequality continued to increase, measures of within-group inequality by education group have diverged. Lemieux (2006) documents that within-group inequality increased for college graduates and postgraduates, remained constant for workers with high school diplomas and some college education, and decreased for high school dropouts. As we argue below, these patterns of within-group inequality are difficult to explain in standard competitive models of the labor market. However we show that a simple frictional model is able to explain these trends and moreover, is consistent with the movements in between-group inequality.

In understanding wage inequality most papers use a competitive framework where wages equal the marginal product of the worker. In these models between-group inequality reflects the price and distribution of observed skills, while within-group inequality reflects the price and distribution of unobserved (to the econometrician) skills. Katz and Murphy (1992), Krueger (1993) and others, use the competitive framework to explain the changing nature of inequality with a focus on skill-biased technical change. But these early papers in the inequality literature were mainly motivated by movements in between-group inequality.

Lemieux (2006) points out that standard accounts of skill-biased technical change have difficulty reconciling the divergence in within-group inequality by group. If skill-biased technical change increases the price of unobserved skill, then within-group wage inequality should increase for all groups. Although this pattern was observed during the 1980s, it breaks down in the post-1990 period. More complicated models with multi-dimensional unobserved skill can explain these facts, but rely upon multiple price movements and/or changes in unobserved skill distributions varying by worker type. Such explanations are inherently difficult to test, and they may not capture how within-group inequality is related to between-group inequality through changes in technology.

There is a second dimension along which competitive models of skill-biased technical change struggle to explain movements in within-group wage inequality. We typically think of changes in the price of unobserved skills in response to skill-biased technical change as being persistent. And if unobserved skills are constant across individuals, or only changing slowly, then the competitive model suggests that the increase in within-group inequality should largely reflect persistent differences in wages

\(^1\)See Katz and Autor (1999), Acemoglu (2002) and Lemieux (2007) for surveys.
across individuals over time. But, the existing empirical evidence suggests the opposite. Gottschalk and Moffitt (1994) and Kambourov and Manovskii (2009) show that much of the increase in within-group inequality has been transitory.

A frictional model of the labor market can address these problems. First, frictional models yield precise implications for how within-group wage distributions change in response to changes in technology. Hence, it is possible to link movements in within- to movements in between-group inequality, a link that is currently underdeveloped. Second, within-group wage inequality in frictional models arises due to luck. Changes in the importance of luck in the labor market may increase the degree of within-group wage inequality. Consistent with empirical evidence, a large component of this increase will be transitory if luck is temporary.

Despite these advantages, the use of frictional models to understand wage inequality has been limited. The work of Burdett and Mortensen (1998) has been successful in explaining wage dispersion - that is, why similar workers are paid differently. However, the assumed homogeneity in worker skills makes the standard Burdett-Mortensen model inappropriate to address issues such as wage inequality where between- and within-group inequality are important features of the data.

To tackle between- and within-group inequality in a unified framework, we develop a frictional labor market model by generalizing the work of Burdett and Mortensen (1998) in Section 2. To extend their model to incorporate heterogeneity in worker productivity we use the concept of skill requirements. We assume that it requires a certain skill level to work for certain firms: the greater the productivity of a firm, the greater the required skill. In all other respects the model remains unchanged. When unemployed, individuals accept wage offers that exceed their reservation value and when employed they accept wage offers that exceed their current wage. On the demand side of the market, firms post wages to maximize profits. We focus on an equilibrium in which more productive firms offer higher wages. Such an equilibrium exists if the productivity differences between firm types are sufficiently large.

Our model is surprisingly tractable. One key equation describes the distribution of workers across firms while a differential equation describes the wages offered by firm type. This allows us to

---

2 Lloyd-Ellis (1999), Violante (2002) and Aghion (2002) study frictional markets and examine (aggregate) within-group inequality. These papers do not address the divergence in within-group inequality by group.

3 Albrecht and Axell (1984), Burdett and Mortensen (1998), Gaumont, Schindler and Wright (2006), and Bontemps, Postel-Vinay and Robin (2000) introduce worker heterogeneity in the form of varying reservation utilities. But, there is little work on production heterogeneity, which we view as more relevant for explaining wage inequality. Van den Berg and Ridder (1998) estimate a model of firm and worker production heterogeneity with segmented markets but their model fails to capture interactions between markets. Finally, Postel-Vinay and Robin (2002) provide a model with worker and firm heterogeneity. We discuss their paper in detail below.
describe the wage distribution conditional upon worker type. Due to skill requirements, workers with low skills are restricted to low productivity jobs while skilled workers are employed across a broader variety of firms. This difference in opportunities generates between-group inequality. Moreover, regardless of skill type, the exact productivity and wage of a worker depends upon luck, hence generating within-group inequality.

Section 3 calibrates the model to match an unemployment rate of 6 per cent and realistic firm productivity distribution. The inclusion of skill requirements helps the model match the observed wage distribution. In particular, the model generates a distribution of wages that is approximately lognormal. In contrast the standard Burdett-Mortensen model without skill requirements struggles to generate realistic wage distributions.

Section 4 examines movements in between- and within-group inequality in the post-1987 period. The salient facts are an increase in between-group inequality with a simultaneous divergence in within-group inequality across groups. As noted above, these trends are difficult to explain in the context of a competitive labor market. We show, however, that in a frictional labor market, skill-biased technical change is a natural mechanism for explaining both between- and within-group inequality. We use the Current Population Survey (CPS) to measure changes in log wage percentiles of the wage distribution. We rely upon the close connection between wages and productivity in the model to extract the implied productivity changes. Our results suggest productivity has increased among the most productive firms, decreased among the mid-range firms, and remained roughly constant for the least productive firms. This “hollowing out” of the productivity distribution (i.e., fewer firms near the median and more firms at the extremes) is consistent with the hypothesis that there has been a polarization of the labor market (Autor, Levy and Murnane (2003)).

Using the calibrated productivity changes, we then ask how within-group inequality responds to skill-biased technological change. We find that within-group inequality increases for the most skilled workers and decreases for the least skilled workers. This complicated outcome arises because different workers are employed across different types of firms. Polarization compresses the productivity and wage offers of low to medium productivity firms. Low skilled workers are employed exclusively at these firms and thus within-group inequality decreases for these workers. On the other hand, skilled workers are employed across both low and high productivity firms and an increase in skill-bias increases the variance of wages offered by more productive firms. This leads to greater within-group inequality among skilled workers capable of working at a wide variety of firms. To the best of our knowledge, our model is the first to be able to simultaneously explain (i) the divergence in within-group inequality by group, as documented by Lemieux (2006), and (ii) changes in within-group inequality driven by transitory shocks.
Several related papers emphasize the role of frictions in explaining wage inequality, notably Shi (2002) and Albrecht and Vroman (2002). Both papers deal with two-sided heterogeneity and examine the response of wages to productivity change but the underlying assumptions differ substantially. These models generate within-group inequality only for some worker types, and thus do not capture differential movements in within-group inequality across groups. Also related are papers by Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay and Robin (2006). They study how worker production heterogeneity contributes to wage dispersion when wages are determined by open bidding between firms. Our paper retains the original Burdett-Mortensen assumption of wage posting, which seems reasonable if potential employers have imperfect information regarding the quality and employment status of a worker. From an empirical perspective, both methods of wage determination are relevant (Hall and Krueger (2008)). Furthermore, the production side of our model features skill requirements, while theirs heterogeneity in efficiency units. Our assumption has the advantage of generating unemployment differences across worker types.

2 A Search Model with Skill Requirements

2.1 Model Setup

The economy consists of a continuum of risk neutral, utility maximizing workers and a continuum of profit maximizing firms. There is an equal and fixed mass of workers and firms in the economy, normalized to one. There is two-sided heterogeneity in the labor market with firms differing in terms of productivity and workers differing in skill levels. Let $p \in [0, 1]$ denote the type of a firm and $P(p)$ the corresponding atomless, strictly increasing, continuous distribution function of firm types. Let $v \in [0, 1]$ denote the type of a worker and $V(v)$ the corresponding atomless, strictly increasing, continuous distribution function of worker types.

Workers search for an acceptable wage while unemployed and for a better wage while on the job. Time is continuous and each worker, employed or unemployed, meet firms with a Poisson arrival rate of $\lambda = 1$, with the exact firm type drawn from $P$. Job matches are destroyed at an exogenous rate $\delta$.

Skill requirements take the following form: a worker with type $v$ can only form a match with a firm of type $p$ if $v \geq p$. If a match is formed, output is divided between worker and firm by a

\[ \text{In detail, Shi (2002) uses a directed search environment while Albecht and Vroman (2002) use a random matching framework with wage bargaining. These models generate wage distributions with three possible realizations. In this respect, the continuous wage distribution generated by our model is a significant advance.} \]
wage rate $w$ that the firm posts prior to meeting a worker. More precisely, if $v \geq p$ then the wage offer is revealed and the worker decides whether to accept or reject the posted wage $w$. If the job offer is accepted, then the flow productivity of the match is $R(p)$ which is a strictly increasing, continuously differentiable function and $R(0) \geq 0$. We will refer to function $R$ as the revenue function, and will interpret this value as the productivity of the firm as captured by the marginal product of a qualified worker. Note that $R(p)$ is independent of worker type but the set of workers capable of working for a firm depends upon the level of productivity. We assume the value of leisure is equal to an exogenous value $z$ for all worker types and that workers and firms discount at rate $r$, which for simplicity we set equal to zero.

2.2 Equilibrium and Analysis

Our equilibrium concept is a stationary equilibrium in which the wage offered by a firm of type $p$, $w(p)$, is strictly increasing in $p$. To define such an equilibrium, let $T_v(p)$ denote the steady-state probability that a worker with type $v$ is either unemployed or works for a firm of type $p$ or less.

**Definition 1.** A stationary monotone equilibrium is (i) a distribution of workers across firms, $T_v(p)$, and (ii) wages offered by different firm types, $w(p)$, such that:

- $T_v(p)$, the distribution of workers across firms over time is constant;
- firms maximize profits;
- workers maximize utility;
- wages are strictly increasing in firm type.

2.3 Distribution of Workers Across Firms

To characterize an equilibrium, we start with the behavior of the workers. Since the rate of matching is independent of job status, unemployed workers will accept job offers that post a wage in excess of their value of leisure, $z$. Employed workers will accept wage offers that exceed their current wage and reject other offers.\(^5\) Now, we are ready to describe the steady-state distribution of workers across firms. Since we consider equilibrium in which $w(p)$ is strictly increasing, the following holds for function $T_v(p)$:

\[ T_v(p) = Pr(wage \leq w(p)|type = v). \]

\(^5\)A formal discussion of the equilibrium behavior of workers is provided in the original Burdett and Mortensen (1998) and the interested reader is referred to their paper for full details.
Since unemployed workers accept all offers and employed workers transit to firms with higher types (they offer higher wages in a monotone equilibrium), the law of motion is

\[ \dot{T}_v(p) = (1 - T_v(p))\delta - T_v(p)(P(v) - P(p)), \]

when \( v \geq p \) and \( \dot{T}_v(p) = 0 \) otherwise. The first term is the flow into unemployment of workers of type \( v \) receiving a wage greater than \( w(p) \), while the second term is the flow of type \( v \) workers from wages below \( w(p) \) to wages above. In a stationary monotone equilibrium, it follows that

\[ T_v(p) = \frac{\delta}{\delta + P(v) - P(p)} \quad \text{(1)} \]

if \( v \geq p \) and \( T_v(p) = 1 \) if \( v < p \) since a worker of type \( v \) will never obtain an offer above \( w(v) \) which is less than \( w(p) \). This is the steady state distribution of workers across firm types and follows from the fact that there is always constrained efficient hiring.

Note that the probability of being unemployed for a type \( v \) worker, \( u(v) \) is

\[ u(v) = T_v(0) = \frac{\delta}{\delta + P(v)}. \]

Consequently, more skilled workers have lower unemployment rates, consistent with empirical evidence.\(^6\) The total level of unemployment in the economy is,

\[ U = \int_0^1 \frac{\delta V'(v)}{\delta + P(v)} \, dv. \]

2.4 Wage Setting

We now focus on the wage-setting decision of a firm and begin by describing the amount of labor that a firm will be able to employ in a monotone equilibrium. To simplify exposition instead of characterizing the profit of a firm from setting a wage level \( w \), we make the decision variable of a firm the type he pretends to be, \( \hat{p} \), by setting wage \( w(\hat{p}) \). Since \( w \) is strictly monotone in firm type the two formulations are equivalent. A firm of type \( p \) will be able to hire workers of type \( v \in [p, 1] \).

In the steady state of a monotone equilibrium, the distribution of workers across firm types is given by \( T_v(p) \) from equation (1). Hence, conditional upon meeting an arbitrary worker, a firm of type \( p \) that offers a wage of \( w(\hat{p}) \), will hire a worker with probability

\[ \int_p^1 \frac{V'(v)}{\delta + \max\{P(v) - P(\hat{p}), 0\}} \, dv. \]

\(^6\)This is an important difference from the work of Postel-Vinay and Robin (2002). Our framework features differential unemployment rates by skill while in their model, unemployment rates are independent of type.
Conditional upon accepting employment, the expected time that a worker of type \( v \) remains with a firm offering a wage of \( w(\hat{p}) \) depends upon the separation rate. Job matches are exogenously destroyed at a rate of \( \delta \) and endogenously destroyed when a worker receives a better offer which occurs at a rate of \( P(v) - P(\hat{p}) \) if \( v > \hat{p} \) and zero otherwise. This implies that conditional upon hiring, the expected employment duration of a type \( v \) at a firm that offers a wage of \( w(\hat{p}) \) is

\[
\frac{1}{\delta + \max\{P(v) - P(\hat{p}), 0\}}.
\]

Combining these implies that a firm of type \( p \) that offers a wage \( w(\hat{p}) \) in a monotone equilibrium will expect to hire workers for the time length conditional upon meeting,

\[
M(p, \hat{p}) = \begin{cases} 
\int_{p}^{1} \frac{V'(v)\delta}{(\delta + P(v) - P(\hat{p}))^2} dv & \text{for } p > \hat{p} \\
\int_{\hat{p}}^{1} \frac{V'(v)\delta}{(\delta + P(v) - P(\hat{p}))^2} dv + \int_{\hat{p}}^{p} \frac{V'(v)\delta}{\delta^2} dv & \text{for } p < \hat{p}.
\end{cases}
\]

Using this, the profit function of a firm with type \( p \) who offers a wage \( w(\hat{p}) \) can be written as,

\[
\pi(p, \hat{p}) = (R(p) - w(\hat{p}))M(p, \hat{p}),
\]

where we use the assumption that the discount rate of future profits is zero. The first order condition implies\(^7\) that in equilibrium

\[
(R(p) - w(p))M^{(2)}(p, p) - w'(p)M(p, p) = 0
\]

where \( M^{(2)}(p, \hat{p}) \) is the derivative of \( M \) with respect to the second variable. Equivalently,

\[
w'(p) = \frac{(R(p) - w(p))M^{(2)}(p, p)}{M(p, p)} \tag{2}
\]

As in the standard Burdett-Mortensen model, search frictions imply that firms have some market power and set wages in a monopolistically competitive market.\(^8\) In particular, wages are set to balance the ability to attract workers with the desire to retain as much surplus as possible from match formation. Changes in the wages offered by firms arise due to changes in the underlying distributions of workers or firms, or via changes in the revenue function.

If a monotone equilibrium exists, profit maximization implies \( w(p) \) must satisfy the differential equation implied by the first order condition in equation (2). In the special case in which \( z = \)

\(^7\)Since \( w \) is assumed to be strictly increasing, it is almost everywhere differentiable. The differential analysis then characterizes \( w' \) almost everywhere and pins down \( w \) for all values using that \( w \) is continuous and that \( w(0) = z \) must hold.

\(^8\)Perhaps the easiest way to see this is to note that the first order condition for wages can be rewritten as

\[
\frac{R(p) - w(p)}{w(p)} = \frac{w'(p)}{w(p)} \frac{M(p, p)}{M^{(2)}(p, p)}
\]

which is the standard mark-up pricing formula associated with monopolistic competition.
$R(0) = 0$ it is also known that $w(0) = 0$. Defining $\tau(p) = \frac{M^{(2)}(p,p)}{M(p,p)}$, the solution to this first order linear differential equation with accompanying boundary condition is given by the following,

$$w(p) = \frac{\int_0^p R(x)\tau(x)Exp\left(\int_0^x \tau(u)du\right)dx}{Exp\left(\int_0^p \tau(u)du\right)}.$$  \hspace{1cm} (3)

In more general situations when $R(0) \geq z \geq 0$, wages will still satisfy the differential equation (2) with the boundary condition $w(0) = z$. As in the standard Burdett-Mortensen model, the least productive firms will offer wages high enough to attract workers but will not try to retain them in the face of competing offers. Note that the differential equation is Lipschitz continuous in $w$ so the above initial value problem has a unique solution for any value of $z \geq 0$.

To analyze whether the above necessary conditions for the wage function $w$ are also sufficient and to have closed form solutions, it is assumed from now that $V(x) = P(x)$, that is the distributions of worker skill levels coincides with the distribution of firm skill requirements. Then, one may assume that $V(x) = P(x) = x$, without (a further) loss of generality.\(^9\) Under this simplification, the Appendix proves the following proposition:

**Proposition 1.** *When types are normalized such that $V(x) = P(x) = x$ and $R(0) = 0$, a sufficient condition for the existence of a unique monotone equilibrium is that $R(p)$ is a (weakly) convex function.*

In some cases a monotone equilibrium does not exist. The reason is that with skill requirements different firms face different sets of potential employees when making wage offers. Hence, the elasticity of labor with respect to the wage varies across firms. In particular, more productive firms may face a smaller elasticity, inducing them to offer lower wages. This effect is absent from models where the form of the ex post competition is Bertrand (as in the work of Postel-Vinay and Robin (2002)) or in models where workers are homogenous (as in the original work of Burdett

\(^9\)To see this, start with a model with general (but equal) distribution functions. Then the distribution function of the maximum production a worker is capable of is

$$Pr(R(v) \leq y) = V(R^{-1}(y)).$$

Then consider a model where the distributions are uniform and the production function is

$$\tilde{R}(p) = R(V^{-1}(p)).$$

Then

$$Pr(\tilde{R}(v) \leq y) = \tilde{R}^{-1}(y) = V(R^{-1}(y)),$$

which implies that the two specifications describe the same model just rescaling variables. In the two models related as above firms with the same productivity post the same wages in equilibrium. Moreover, a worker capable of working for a given productivity firm (and for less productive ones) would have the same wage histories in the two specifications. Therefore, the two models are equivalent and performing this transformation is payoff irrelevant and would leave all our results exactly intact.

9
and Mortensen (1998)). As a consequence, in those models a more productive firm always offer
more attractive wages. In our model this may not be the case because of the different elasticities
faced by different firms. To assure that incentives remain monotone, the above proposition assumes
convexity of $R(p)$, which ensures that the productivity differences between firms are large enough
so that a monotone equilibrium exists.\footnote{This strong sufficient condition for existence is not always meet when we calibrate our model. Therefore, we check numerically profit maximization holds.}

\section{The Shape of the Wage Distribution}

\subsection{Calibration}

The model is challenging to calibrate since the primitives include a distribution of worker types,
a distribution of firm types, and a revenue function. Inherently, the distributions of worker and
firm types are difficult to observe as they relate to the way skill requirements limit the ability of
particular workers to match with certain firms. Hence, our calibration strategy starts from the
basic simplifying assumption that worker and firm types are distributed uniformly. We then select
a revenue function that generates an observed distribution of firm productivity comparable to the
distribution observed by Hsieh and Klenow (Forthcoming).

To begin, the value of $\delta$ is set to to match the 6 percent unemployment rate observed over the
period from 1987-1997, normalizing (as before) $\lambda$ to 1. The endpoints of this period cover the
points in time at which Hsieh and Klenow measure the distribution of productivity. The aggregate
level of unemployment is,

$$U = \int_0^1 \frac{\delta}{\delta + v} dv = \delta \log \left(1 + \frac{1}{\delta}\right),$$

which suggests a value of $\delta = 0.014$ is required generate an unemployment rate of 6 per cent.\footnote{This value of $\delta$ is lower than in many other search models of the labor market; but recall, equilibrium outcomes are determined by the ratio of $\delta/\lambda$. Shimer (2005) sets a value of $\delta = 0.1$ and a mean value of the job finding rate of 1.35. In contrast, the aggregate job finding rate under this calibration is 0.22 which implies a much lower value of $\delta$ is necessary to match aggregate unemployment.}

To choose a reasonable functional form for the revenue function we appeal to the recent work of
Hsieh and Klenow (Forthcoming). They use the Census of Manufacturers collected by the US
Bureau of Census to extract a distribution of firm productivity and find that it is approximately
log-normally distributed. Note that given an underlying uniform distribution of worker and firm
types, there is an implied distribution of actively producing firms in equilibrium. Define $G(p)$ to be
the cumulative distribution function of firm types that are matched with workers in equilibrium.
That is,

\[ G(p) = Pr(firm\ type \leq p|matched) . \]

The appendix shows that in stationary equilibrium,

\[ G(p) = \frac{p - \delta \log(1 + \delta) + \delta \log(1 - p + \delta)}{1 - \delta \log \left(1 + \frac{\delta}{2}\right)} . \tag{4} \]

Given this equilibrium CDF of firm types, we select a revenue function to generate a distribution of firm productivity that matches the lognormal distribution displayed by the data. To do so requires a revenue function of the following form,

\[ R(p) = \exp(\sigma \sqrt{2}\text{erf}^{-1}(2G(p) - 1) + \mu) + d \]

where \( \text{erf}^{-1} \) is the inverse error function.\(^{12}\) In this parameterization, \( \mu \) and \( \sigma \) are the mean and standard deviation of the underlying normal distribution and \( d \) is a parameter that allows us to shift the distribution. The introduction of this shift parameter lets us consider situations in which the minimum productivity is non-zero and enables firms to make positive profits as long as \( d \) exceeds the reservation utility of workers. Hence it also allows us to consider realistic calibrations with a non-zero flow value of leisure.

To complete the calibration of the revenue function, we need to choose values for \( \mu, \sigma \) and \( d \). To do so we use three pieces of information. First, we set \( d \), the minimum productivity level associated with a firm to equal 3.35, which is the minimum wage in 1987. Intuitively, firms with productivity below the minimum wage will not survive in a long-run equilibrium. Second, Hsieh and Klenow (Forthcoming) find that the standard deviation of log productivity in 1987 in their sample of manufacturing firms is equal to 0.41. However, their procedure does not allow us to extract the average level of productivity since they measure the dispersion of detrended productivity. To proceed, we take the OECD estimate of for US labor productivity in 2007 in current prices. We deflate by the CPI and use the BLS index of business output per hour to extract an estimate for 1987 US labor productivity. Finally, \( R(p) \) is interpreted as the marginal product of labor so we adjust by a labor share factor of 0.66. This yields enough information to pin down the parameters of the revenue function.

This calibration strategy may understate productivity dispersion. First, the data presented by Hsieh and Klenow captures the dispersion in average productivity between plants. In addition,

\[^{12}\text{To see why this revenue function generates a distribution of productivity that is lognormal, note that if we select a revenue function of } R(p) = G(p), \text{ the distribution of firm productivity in equilibrium will be uniform over the interval } [0,1]. \text{ To transform a random variable, say } x, \text{ with a uniform distribution over the interval } [0,1], \text{ to a log-normal distribution with parameters } (\mu, \sigma) \text{ the appropriate transformation requires that } f(x) = \exp(\sigma \sqrt{2}\text{erf}^{-1}(2x - 1) + \mu).\]
there may be productivity dispersion among individual workers within plants that is important for understanding wage inequality but is not captured by the Census data. Second, sample selection is an issue: the Census of Manufacturers only covers manufacturing firms. Examining the UK, Faggio, Salvanes, and Van Reenen (2007) note that productivity in manufacturing firms is more compressed than in the aggregate since manufacturing is a declining industry.

The final parameter to calibrate is $z$, the value of leisure. The literature examining business cycle volatility in the labor market has reached little agreement on the value of this parameter. Shimer (2005) uses a relatively low value while Hagedorn and Manovskii (2008) argue for a high value. For our baseline simulations we set $z = 2$, which is about 60 per cent of the minimum wage. Our results are not sensitive to this assumption.\(^\text{13}\) The key parameters are summarized in Table 1.

### 3.2 Distribution of wages

Given our calibration, we numerically approximate the model and compute the equilibrium wage distribution. The implied distribution of the model with skill requirements is compared to the standard Burdett-Mortensen model and to the data. We find that the introduction of skill requirements generates large changes in wage setting behavior and these changes help the model match the data when there is a reasonable underlying firm productivity distribution.

Figure 1 shows the wage as a function of productivity in our model with skill requirements and in the standard model calibrated to match a similar unemployment rate and an identical productivity distribution. The presence of skill requirements generates an equilibrium in which wage offers nearly equal productivity for the bulk of firms. This close connection between $w(p)$ and $R(p)$ is a key feature we comment on below. In contrast, the standard Burdett-Mortensen model features a much larger deviation between productivity and wages. Formally, this difference arises due to the

\(^{13}\)Varying $z$ over the interval $[0, 3.35]$ only has a small impact upon the wages offered by the majority of firms. Since the calibrated level of frictions ($\delta$) is small, for most firms the differences between productivity and offered wage is small regardless of the reservation utility of workers. Altering the value of $z$ does have an impact upon the wages offered by the least productive firms but these differences narrow quickly. For example, the wages offered by a firm of type $p = 0.05$ when $z = 3.35$ is less than two per cent larger than the wage offered in the case in which $z = 0$.  

\begin{table}[h]
\centering
\begin{tabular}{llll}
\hline
Parameter & Concept & Value & Calibrated to match \\
\hline
$\mu$ & mean of log-normal & 2.03 & Average labor productivity (OECD) \\
$\sigma$ & variance of log-normal & 0.6 & Distribution in Hsieh-Klenow \\
$\delta$ & rate of job destruction & 0.014 & 6 per cent unemployment \\
$\gamma$ & shift parameter & 3.35 & Minimum wage \\
$z$ & value of leisure & 2 & 60 per cent of minimum wage \\
\hline
\end{tabular}
\caption{Summary of Calibrated Values}
\end{table}
differing first order conditions associated with wage setting:

With skill requirements: \[ w'(p) = (R(p) - w(p)) \left( \frac{2}{\delta + 1 - p} \right) \left( \frac{2\delta + 1 - p}{2\delta} \right) \]

Without skill requirements: \[ w'(p) = (R(p) - w(p)) \left( \frac{2}{\delta + 1 - p} \right) \]

where the first condition is associated with skill requirements and the second exists in a model without skill requirements.\(^{14}\) In both calibrated models, \( \delta \) will be small to match low aggregate unemployment. With skill requirements this implies a high value of \( w'(p) \) unless \( R(p) \approx w(p) \). In contrast, in the standard model even when \( \delta \) is small \( w'(p) \) will only become large as \( p \to 1 \).

The intuition behind this result is related to how the equilibrium labor employed by a firm responds to wage increases in both models. In the standard Burdett-Mortensen model (with firm heterogeneity), the combination of on-the-job search and the absence of skill requirements generates a convex distribution of workers over firm types. That is, the distribution of workers is highly skewed to be matched with high productivity firms. Hence, for most firms, an increase in wage attracts relatively few workers and these workers are retained for a short period of time. This limits the incentive of wage increases for most firms in the standard model. In contrast, with skill requirements the distribution of worker types across firms is concave and an increase in wage leads to a relatively large employment effect. To see why, consider a firm of type \( p < 1 \). By increasing wage offers from \( w(p) \) to \( w(p + \epsilon) \) for some small \( \epsilon > 0 \) the firm is able to attract and retain all workers of type \( v \in [p, p + \epsilon] \) who are unable to accept offers from firms of type greater than \( p + \epsilon \). This leads to

\(^{14}\)The first equation is derived in the Appendix, the second equation is easily derived from equation (3.17) of Mortensen (2003), with \( \lambda = 1, \Gamma(p) = p \) and a generalized \( R(p) \).
a large employment response to wage increases and thus firms have an incentive to increase their wage offers if the wage is significantly lower than the marginal product of the worker. In our first order condition (the first equation above) this effect shows up as a large value for $w'(p)$ if $R - w$ is significantly positive.

A second observation is that with skill requirements while wages and revenues are similar for most firm types there are two significant differences: firms with relatively low types and with relatively high types offer significantly lower wages than their revenue. Intuitively, the discrepancies at the two tails arise due to limited competition. By construction, the productivity distribution (log-normal) has relatively few firms in the tails competing for workers. This lack of competition generates a gap between productivity and the wages offered by firms in both the left- and right-tail of the distribution in equilibrium. In the right-tail the workers have many offers, but only a few from very high productivity firms and thus firms with very high productivity are able to extract rents from the workers. In the left-tail firms try to hire workers with low type who can only work for a few firms and thus they are willing to accept low wages.

Figure 2 shows the probability density function of productivity and for wages, with and without skill requirements. By construction, productivity is distributed according to a shifted log-normal distribution. If $w(p) = R(p)$ for all $p$, then wages would be identically distributed. Due to frictions, wages deviate from productivity and hence the distribution of wages deviates from lognormality. With skill requirements, since the difference between $w(p)$ and $R(p)$ is small for most firms, the deviation between wage and productivity distributions is also relatively small. Figure 2 also shows
the distribution of wages in the standard Burdett-Mortensen model without skill requirements. This calibration generates an implied distribution of wages that deviates significantly from lognormality. In particular, the standard Burdett-Mortensen model can not generate the long right-tail that is characteristic of lognormal distributions when the underlying productivity distribution is lognormal.\(^{15}\) This leaves the standard model ill-suited to explain the right-tail of the wage distribution, where much of the recent change in wage inequality has been focused.\(^{16}\)

With skill requirements, one area in which the productivity and wage distributions differ markedly is at bottom where our wage distribution has a large mass relative to the lognormal productivity distribution. This is consistent with Battistin, Blundell and Lewbel (2007). They note that although wage distributions are approximately lognormal, they display a large mass of workers at the bottom-end of the wage distribution (relative to lognormality). As shown in Figure refwrpdf, our model generates a large mass of workers at the bottom-end of the wage distribution relative to the productivity distribution. This arises for two reasons. First, the minimum productivity is exogenously set at a value of \(d = 3.35\) while the minimum wage offered will equal the value of \(z = 2.\) Second, even when \(z\) is closer to \(R(0)\) the lack of competition discussed above implies that firms will offer \(w(p)\) well below \(R(p)\) for low values of \(p.\) Consequently, the fact that the empirical wage distributions feature a relatively large mass of workers at the bottom-end may arise due to imperfectly competitive labor markets.

We also evaluate the standard model and the model with skill requirements by comparing moments of the model generated wage distribution to the actual 1987 wage distribution. The results are

\(^{15}\)Section 4.3.3 of Mortensen (2003) shows to generate the right-tail of a lognormal wage distribution the standard model requires unrealistically high productivity. The result discussed here can be seen as the other side of the same coin; reasonable productivity distributions can not generate the right-tail of wage distributions in the standard model.

\(^{16}\)To see this note that the maximum hourly wage offered in the calibration of the standard model is 15.15. This corresponds to about the 90-th percentile of the wage distribution in the MORG database. In contrast, with skill requirements the maximum wage offered is about four times larger.
shown in Table 2. A few key issues are worth highlighting. First, both models overstate the average level of wages. This suggests that either the level of productivity is overestimated in our calibration, or that both models overstate the wage to productivity ratio. Second, although the model with skill requirements overestimates the mean of wages by a greater amount it performs significantly better on all other measures (standard deviation, skewness, and kurtosis) than the standard model.

The model also has implications regarding the magnitude of wage inequality. Computing the standard deviation of log wages yields a value of 0.43 which is marginally larger than the underlying inequality in productivity (0.41). In a frictionless environment the model converges to a perfectly competitive equilibrium with \( w(p) = R(p) \); in this case, measures of wage and productivity inequality converge. In a world with frictions they differ for two reasons. First, firms offer wages below productivity as they maintain some of the match surplus. This results in compression of the wage structure and less wage than productivity inequality. Second, when \( z < d \), the fact that the value of leisure is lower than the minimum productivity level implies that wages will start out at a lower level and may lead to greater wage than productivity inequality. In our calibration, the second effect is stronger yielding a marginally larger wage inequality than productivity inequality.

### 3.3 Within-group wage inequality

A key feature of the model is that it generates endogenous within-group inequality with a continuum of worker groups. Due to search frictions, workers with identical ability experience different fortunes in the labor market and receive different wages. This within-group inequality when groups are defined by type, is transitory in nature: over an infinite horizon, the average wage a worker earns converges to the group (type) average. The transitory nature of within-group inequality is consistent with the data as we argued in the introduction.

It is also interesting to analyze whether within-group inequality (as defined for our narrow groups) is consistent with empirical observations. The standard deviation of log wages for narrowly defined worker types is shown in Figure 3. Relative to the data, the model understates the degree of within-group wage inequality. For example, when groups are defined according to education, experience and gender, the standard deviation of log wages varies from roughly 0.15 to 0.30, depending upon group.\(^{17}\) In our calibration, the standard deviation of log wages for the majority of workers lies below 0.15. This result is not surprising; our model assumes that types vary continuously and are perfectly observable while empirical studies measure groups at a very coarse level.

\(^{17}\)For exact figures see Table 1 of Lemieux (2006).
A noticeable feature is that the magnitude of within-group inequality varies in a non-monotonic fashion. The most skilled workers are subject to the greatest amount of within-group inequality but the model implies that some low skilled workers experience greater within-group inequality than workers in the middle of the skill distribution. The possibility of nonmonotonicity arises, because there are two competing effects. First, workers with higher types have a larger upper bound for their wages, while the lower bound is the same for all types. This effect implies a higher standard deviation in (log) wages for high types. Second, the calibrated revenue function $R$ is very steep for low values of $p$, because there are only a few firms with very low productivity if firm productivity is distributed lognormally. The close connection between wages and revenues imply that $w$ increases rapidly in $p$ when $p$ is low. Hence, low skilled workers employed at very low productivity firms will face large wage dispersion, since $w$ is very spread out for the small set of firms they work for. This second effect implies that wage dispersion is high for low type workers, but then decreases. The sum of the two effects is a non-monotone wage inequality in the type of the workers. Although this is an interesting theoretical prediction, we would not necessarily expect it to be confirmed in the data, where groups are imperfectly measured and are likely to be broad aggregates of worker types in our model.

It is interesting to point out that the calibrated level of search frictions are relatively low, and thus firms are offering wages close to the competitive wage, but at the same time a significant degree of within-group inequality exists. Recall, however, that there are two dimensions along which this model varies from a competitive economy. First, firms offer wages that deviate from competitive outcomes due to matching frictions and second, (employed) workers of a given type are distributed across a variety of firm types. It is this second deviation from the competitive model which is
responsible for generating within-group inequality in our model.

4  Skill-Biased Technical Change and Wage Inequality

We now examine the comparative statics of skill-biased technical change. Special attention is devoted to understanding the joint behavior of between- and within-group inequality. We concentrate on a period where the between-group inequality increases and within-group inequality diverges by group.\textsuperscript{18} The key features of this period are that the wage premium for college education increased while wage differentials associated with low levels of education remained relatively constant or decreased. Simultaneously, there is a divergence in within-group inequality across education groups.

4.1  Review of changes in wage inequality

We begin reviewing trends in between-group wage inequality since 1980. Table 3 summarizes wage differentials by education relative to high school graduates over time. These differentials are constructed from the Merged Outgoing Rotation Groups (MORG) of the CPS by using a log wage regression with education levels represented by dummy variables and controls included for experience, marital status and race.\textsuperscript{19} A number of features are apparent. Firstly, the changes in wage differentials have been large; the male college wage premium (relative to high school graduates) has increased from 0.29 in 1980 to 0.49 in 2000. It is also notable that much of the rise in inequality is concentrated in the 1980s. Of the 0.2 increase in log wage points of the male college wage premium, about 0.15, or 75 per cent, occurred during the 1980s. The changes in between-group wage inequality in the 1990s, in addition to being smaller, were less pervasive and focused more on skilled workers. Although the education wage premium increased for both male and female college graduates and post-graduates in the 1990s, there was a decrease in the wage premium associated with “some college” education. Finally, the magnitude of the wage differential between high school dropouts and high school graduates remained stable for females and increased marginally for males.

Lemieux (2006) details the changing pattern of within-group wage inequality. His results regarding the evolution of male within-group inequality are reproduced here in Figure 4 and the pattern

\textsuperscript{18}Periods where within- and between-group measures move in the same direction for all groups are easy enough to explain in many models, including ours. Prior to 1987 both within- and between-group measures increase. Then our model with a proportional increase in $R(p)$ is sufficient to capture these movements qualitatively.

\textsuperscript{19}The underlying data is adjusted to take account of top-coding and the sample is restricted to exclude outliers. See Lemieux (2006) for specific details of the adjustment process.
of changing female within-group wage inequality is broadly similar. After a period of stability during the 1970s, within-group inequality increased rapidly for all groups during the early 1980s. From the late 1980s a new pattern emerges; within-group inequality for skilled workers (college or postgraduate education) has steadily increased while it has generally remained constant or decreased for less skilled workers (high school dropouts, high school graduates and some college).\footnote{Although Lemieux also examines the evolution of total residual inequality in his paper, we focus on within-group inequality on a group-by-group basis since this is a more informative measure of the evolution of wage inequality.}

Autor, Katz, and Kearney (2006) propose a model based upon varying occupational choices leading to variations in unobserved distribution of skill in different groups. Their story relies on individuals making choices that alter the distribution of unobserved skills of different groups. Although their model can generate movements in within-group inequality consistent with Lemieux (2006), it also suggests, counterfactually, that these changes should be permanent in nature. Moreover, as documented by Shimer (1998) there have been vast changes in educational attainment and hence changes in the composition of educational groups, going back at least until the 1970s. This raises a problematic question for Autor, Katz, and Kearney. Why is it that we observe this differential movement in within-group inequality only since the late 1980s despite large changes in educational choices and hence, presumably group composition as early as the 1970s?

Our model, on the other hand, moves away from thinking about within-group inequality due to unobserved skill and focuses attention on luck in a frictional labor market. To examine whether, in a frictional environment, a change in skill-biased technical change is capable of generating the observed between- and within-group movements in inequality we conduct a comparative statics exercise to examine how within-group inequality responds to skill-biased technical change.

### 4.2 Skill-biased technical change

Our starting point is the baseline calibration of Section 3. This calibration provides our estimate for the productivity and wage distributions in 1987. To understand how productivity has evolved we rely upon the close connection between wages and productivity implied by the model. Using

---

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High School Dropouts</td>
<td>-0.26</td>
<td>-0.30</td>
<td>-0.32</td>
<td>-0.22</td>
<td>-0.28</td>
<td>-0.28</td>
</tr>
<tr>
<td>Some College</td>
<td>0.09</td>
<td>0.15</td>
<td>0.12</td>
<td>0.12</td>
<td>0.19</td>
<td>0.16</td>
</tr>
<tr>
<td>College Graduates</td>
<td>0.29</td>
<td>0.44</td>
<td>0.49</td>
<td>0.31</td>
<td>0.49</td>
<td>0.53</td>
</tr>
<tr>
<td>Post-graduates</td>
<td>0.37</td>
<td>0.58</td>
<td>0.65</td>
<td>0.50</td>
<td>0.68</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Measures of between-group inequality relative to high school graduates. Controls for experience, race and marital status.
the MORG data we construct the change in log weekly real wages for each percentile between the fifth and ninety-fifth percentile, from 1987-1997.\footnote{Focusing on the fifth to the ninety-fifth percentile reduces issues associated with measurement error and topcoding. The period from 1987-1997 is selected to be comparable with the period over which Hsieh and Klenow examine the distribution of productivity.} The results are displayed in Figure 5 and are consistent with the findings of Autor, Katz, and Kearney (2006) over the slightly different period 1990-2000. We find a non-linear, U-shaped relationship between the change in log wages and the position in the wage distribution. Real wages in the middle of the distribution have stagnated, declining by about 0.05 log wage points. On the other hand, wages at the top end of the distribution have increased by roughly the same amount. Finally, wages at the bottom of the distribution have remained roughly constant, in real terms.

We use this result to pose a basic question: what change in the revenue function would create such a change in the percentiles of the wage distribution? As previously noted, our model implies a close connection between the wages and productivity and suggests that the percentage change in wages received at each percentile must be similar to the percentage change in productivity for the corresponding firm.\footnote{In our calibrated model, for firms between the 10th and the 90th percentile, the $\frac{w(p)}{R(p)}$ ratio is almost constant, varying between 0.963 and 0.986.} This suggests the approximation that the change in log wages offered at a particular percentile is equal to the change in log productivity for the firm offering such a wage.

The above assumption yields a discrete set of implied productivity changes for firms at different positions of the productivity distribution. However, our model features a continuous revenue function. To estimate how the revenue function changes, we take our discrete set of productivity changes implied by wage changes and we use an OLS regression to estimate the change in log pro-
Figure 5: Changes in Log Real Wages by Percentile of Wage Distribution, 1987-1997

Table 4: Summary of Regression Equation (5)

<table>
<thead>
<tr>
<th>Parameter: Coefficient</th>
<th>Standard Error:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>-0.02 (0.004)</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-0.2 (0.019)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.29 (0.019)</td>
</tr>
</tbody>
</table>

R-squared: 0.87

ductivity depending upon the position in the productivity distribution, $p$. The following quadratic specification is used to capture the non-linearities that are clearly observed in the change in log wages,

$$\Delta \log R(p) = \alpha_0 + \alpha_1 p + \alpha_2 p^2 + \epsilon \quad (5)$$

The regression results are provided below in Table 4. Despite its simplicity, the quadratic specification appears to do a good job of capturing the underlying changes in log wages.23 The fitted values from this regression are used to generate a revenue function for 1997. In particular let,

$$R_{1997}(p) = (1 + Q(p)) \cdot R_{1987}(p)$$

where $Q(p)$ is the predicted change in log productivity implied by regression (5), $R_{1987}(p)$ and $R_{1997}(p)$ are the revenue functions for 1987 and 1997, respectively. For our initial revenue function, $R_{1987}(p)$, we use the baseline calibration in Section 3. The $R_{1997}(p)$ revenue function updates the 1987 revenue function, by taking into account the changing nature of log wages. This implied
Figure 6: Changes in log wages at different percentiles: model versus actual data

productivity change increases the standard deviation of log productivity in the model, from 0.41 to 0.44. This is smaller in magnitude but in the same direction as the change in productivity observed in the manufacturing data. Hsieh and Klenow (Forthcoming) find that the standard deviation of log productivity among manufacturing firms increases from 0.41 to 0.49 over the period from 1987-1997.

We can verify that applying this form of skill-biased technical change to our model does a good job of replicating the changes in log wages at different percentiles of the wage distribution. Figure 6 plots for each percentile the change in log wages in the actual data from 1987-1997 against the change in log wages implied by the model as a result of technical change. The close fit simply confirms that the approximating log wage changes with productivity changes and that the quadratic assumption regarding the form of skill-biased technical change are accurate.

Given this change in productivity, we examine how within-group inequality responds. That is, we compute the standard deviation of log wages conditional upon worker type for our different revenue functions, $R_{1987}$ and $R_{1997}$, and the implied change in within-group inequality over time. The results are presented in Figure 7. For the least skilled workers, the change in productivity reduces within-group inequality. In contrast, for more skilled workers the implied productivity change raises within-group inequality. The complicated, non-linear impact of technology upon within-group inequality arises because different workers are employed across different types of firms. The technical change considered here compresses productivity of the least productive firms. Low skilled workers find employment at these firms and thus within-group wage inequality decreases for this set of workers. On the other hand, technical change expands the productivity dispersion of the most...
productive firms. Skilled workers are mostly employed at these firms, leading to greater within-group inequality among skilled workers. Hence, at least in this calibrated example, technological change consistent with movements in the percentiles of log wages tends to increase the average wage of more skilled workers and have a differential impact upon within-group inequality that depends upon skill type.

4.3 Aggregating Worker Types to Educational Groups

The previous analysis examines how between- and within-group inequality respond to a productivity change derived from observed movements in wages. Figure 6 shows that the calibrated changes for each wage percentile are consistent with the data, and thus our model is able to capture changes in between-group inequality numerically well. Figure 7 and the discussion in the previous section then shows that if we calibrate the model to fit changes in between-group inequality, then the model is able to capture qualitative features of changes in within-group inequality based on narrowly defined groups. However, a difficulty in mapping our model to data is that our model features a continuum of worker types while only a finite number of groups are observable in the data. To bring our work closer to the data, it seems natural to introduce aggregate groups into our model economy. In particular, we assume that there are five separate educational groups: high school dropouts (hsd), high school graduates (hsg), some college (sc), college graduates (cg), and postgraduates (pg). Each educational group consists of a mix of different workers which may be represented by an underlying probability distribution of worker types.

An obvious difficulty lies in estimating the underlying probability distribution of worker types for
Figure 8: Implied firm type associated with corresponding percentile in the wage distribution by education group.

We do not attempt such an ambitious task in this paper. Instead, we note that the MORG database allows us to extract the \( i \)-th percentile of the wage distribution received by educational group, \( j \), which we shall call \( w_{i,j} \). We proceed by extracting from the 1987 MORG the wage percentiles by each educational group. Furthermore, our model describes the wage offered by firm type, conditional upon the revenue function. For the wage, \( w_{i,j} \), we use the wage function associated with the 1987 calibration to extract the implied firm type that offers the corresponding wage and define this firm type, \( p_{i,j} \). This procedure essentially extracts the percentiles associated with the distribution of workers over firms for each educational group. The results are presented in Figure 8 which shows, for each educational group, how each percentile of the wage distribution is matched to firm type.

This procedure does not allow us to extract the complete probability distribution of worker types underlying each educational group. Hence, we can not calculate, for example, the standard deviation of log wages associated with an educational group. But it does provide a set of reasonable restrictions that the underlying distribution of education groups should satisfy and furthermore, for a given \( w(p) \) we are able to compute other well-known measures of inequality such as the 90-10 and the 75-25 log wage differentials.

We then ask the following question: How does within-group inequality respond for our aggregate groups when there is a productivity change (as examined in Section 4.2)? In particular, are our inferences drawn from a a continuum of worker types still valid when we consider a setting with aggregate groups that satisfy the above distribution restrictions? We address this question by calculating the actual change in the 90-10 log wage differential for each educational group from
1987-1997 using the MORG data. Then we compare this to the change in the 90-10 log wage differential for the corresponding educational group, implied by the model, in response to the productivity change outlined in Section 4.2.

In investigating the impact of a skill-biased technical change as in Section 4.2, we assume that matching frictions and the underlying worker and firm distributions are unchanged so that the equilibrium distribution of workers across firm types is constant.\textsuperscript{24} All of the change in the wage distribution is driven by changes in the revenue function. This implies the change in the 90-10 log wage differential due to skill-biased technical change can be calculated as follows,

\[
(\log(w_{1997}(p_{90,j})) - \log(w_{1997}(p_{10,j}))) - (\log(w_{1987}(p_{90,j})) - \log(w_{1987}(p_{10,j})))
\]

where \(w_{1987}\) and \(w_{1997}\) are the wage functions corresponding to the \(R_{1987}\) and \(R_{1997}\) revenue functions, respectively.

The results of this exercise are shown in the left-hand panel of Figure 9. There are a num-

\textsuperscript{24} A number of papers examine wage inequality with endogenous entry. (Julien, Kennes, and King (2006) and Cuadras-Morato and Mateos-Planas (2006)). In the context of this model, an endogenous response of vacancies could be added by using a zero profit condition. See Mortensen (2003) for a discussion of endogenizing firm entry. Such an endogenous entry might moderate the wage response, but we expect that the results would not change qualitatively.
ber of notable features. Firstly, the change in the log wage 90-10 differential is consistent with Lemieux’s (2006) measure of changing inequality. The data suggests that within-group inequality has tended to increase for workers with some college, college and postgraduate education. Within-group inequality for high school graduates is roughly unchanged and decreases for high school dropouts. In response to the productivity change outlined in Section 4.2, the model predicts similar results. Within-group inequality for high school dropouts decreases, while staying approximately constant for high school graduates, and increasing for other groups. The absolute magnitude of the changes is slightly underestimated by our model.

We conduct a similar exercise with the 75-25 log wage differential. The results are displayed in the right-hand panel of Figure 9. In this case the direction of movement of within-group inequality is predicted correctly for all groups except for high school graduates. The model underestimates the increase in the 75-25 log wage differential among workers with some college education but does well in predicting the magnitude and direction of changes in inequality for the remaining groups. Although the match between model and data is not perfect, the current analysis suggests that the mechanism that reduced within-group inequality among unskilled workers and increased it for skilled workers is still present when aggregating across worker types in a reasonable fashion.

Implicit in this analysis is the assumption that changes in the wage distribution reflect changes in the revenue function, rather than changes in the underlying distribution of worker or firm types. Potentially, the wage distribution could be affected by either factor. To get a handle upon whether changes in the underlying distributions of worker or firm types by educational group have been important we show how the unemployment rate by education group has evolved over time, relative to the unemployment rate of high school graduates. Although there are large cyclical fluctuations in unemployment, Figure 10 shows the relative unemployment rate of each educational group remains roughly constant over time. This suggests that changes in the underlying distributions of worker or firm types plays only a minor role in explaining within-group inequality.

This example illustrates that recent interpretations of skill-biased technical change are able to generate between-group inequality movements consistent with the data. Going beyond previous work, this paper also shows that skill-biased technical change is also able to explain detailed movements in within-group wage inequality. Crucial to our results is the presence of heterogeneity which allows

---

25 Daly, Jackson and Valetta (2007) undertake a similar exercise using slightly more disaggregated data. Over the period of interest, the results are consistent.

26 A change in the distribution of firm skill requirements would create differential unemployment effects across groups as it is easily seen from our equilibrium conditions. Potentially, a joint change in the composition of education groups and in the distribution of firm skill requirements could keep relative unemployment rates constant. However, as we argued in footnote 9 such changes may not be identified separately, and thus changes in either firm skill requirements or worker types (but not both) have consequences for equilibrium outcomes.

26
A key aspect of our frictional model of wages is that changes in technology have an impact upon between-group inequality but also within-group inequality in a well-defined manner.

5 Conclusion

We extend the Burdett-Mortensen model to include two-sided heterogeneity by incorporating skill requirements in the production process. Since more skilled workers are able to work at more productive firms we generate between-group inequality similar to competitive models of the labor markets. Unlike standard competitive models, frictions in the matching process generate within-group inequality that is a consequence of luck rather than unobserved heterogeneity. Hence we can discuss both within- and between-group inequality in a unified framework.

We propose a calibration to match a reasonable unemployment rate and a shifted lognormal distribution of underlying productivity. Given this calibration, we show that the underlying wage distribution compares favorably to the standard Burdett-Mortensen model and does well in fitting the data. The nature of imperfect competition implies a large mass of workers towards the bottom-end of the wage distribution (relative to lognormality) as observed by Battistin, Blundell and Lewbel (2007). And relative to the standard Burdett-Mortensen model, our model is better able to capture the long right-tail of the wage distribution.

Finally, we use the model to explain movements in inequality in the post-1987 period. During
this period there has been an increase in between-group inequality and a divergence in within-
group inequality depending upon group. Using changes in log wages we extract, using our model, 
estimated changes in log productivity for different firm types. We then consider the impact of 
this implied change in productivity upon within-group inequality. Our main finding is consistent 
with the empirical evidence of Lemieux (2006). In particular, in response to skill-biased technical 
change within-group inequality for skilled workers tends to increase while for unskilled workers there 
is a decrease. Unlike other competing explanations such as Autor, Katz and Kearney (2006), our 
model is also consistent with the finding that much of the increase in within-group wage inequality 
is transitory in nature (Gottschalk and Moffitt (1994) and Kambourov and Manovskii (2009)). We 
know of no other model capable of explaining this combined set of facts.

6 Appendix

6.1 Proof of Proposition 1

The proposed monotone equilibrium is characterized by a function $w(p)$, that solves the differential 
equation (2) with boundary condition $w(0) = 0$. This section shows that when $V(x) = P(x) = x$ 
and $R(p)$ is convex, that a monotone equilibrium exists.

All we need to do is to check that global second order conditions are satisfied for this candidate 
equilibrium. For checking the second order conditions it is sufficient to prove that

$$
\pi^{(2)}(p, \hat{p}) < 0 \quad \text{if } p < \hat{p}
$$

$$
\pi^{(2)}(p, \hat{p}) > 0 \quad \text{if } p > \hat{p}.
$$

For example, the second condition means that a high type always has a marginal incentive to 
increase a bid from $w(\hat{p})$, which is sufficient to rule out that there is an incentive to deviate 
downwards. The first condition works in the opposite direction.

Let us check the first condition. Then,

$$
\pi^{(2)}(p, \hat{p}) = -w'(\hat{p})M(p, \hat{p}) + (R(p) - w(\hat{p}))M^{(2)}(p, \hat{p})
$$

$$
= (R(p) - R(\hat{p}))M^{(2)}(p, \hat{p}) + w'(\hat{p}) \left( \frac{M(\hat{p}, \hat{p})}{M^{(2)}(\hat{p}, \hat{p})} M^{(2)}(p, \hat{p}) - M(p, \hat{p}) \right)
$$

using the first order condition to eliminate $w(\hat{p})$. 28
Then for $\hat{p} > p$, 

$$M(p, \hat{p}) = \int_{p}^{\hat{p}} \frac{V'(v)}{\delta} \, dv + \int_{\hat{p}}^{1} \frac{\delta V'(v)}{(\delta + P(v) - P(\hat{p}))^2} \, dv > M(\hat{p}, \hat{p})$$

and 

$$M^{(2)}(p, \hat{p}) = \int_{p}^{1} \frac{2\delta P'(\hat{p}) V'(v)}{(\delta + P(v) - P(\hat{p}))^3} \, dv + \int_{\hat{p}}^{1} 0 \, dv = M^{(2)}(\hat{p}, \hat{p})$$

implies,

$$\left( \frac{M(\hat{p}, \hat{p})}{M^{(2)}(\hat{p}, \hat{p})} M^{(2)}(p, \hat{p}) - M(p, \hat{p}) \right) < 0$$

Note that $w'(\hat{p}) > 0, M^{(2)}(p, \hat{p}) > 0$ and $(R(p) - R(\hat{p})) < 0$ implies $\pi^{(2)}(p, \hat{p}) < 0$

Now, consider the second condition,

$$\pi^{(2)}(p, \hat{p}) = (R(p) - R(\hat{p})) M^{(2)}(p, \hat{p}) + w'(\hat{p}) \left( \frac{M(\hat{p}, \hat{p})}{M^{(2)}(\hat{p}, \hat{p})} M^{(2)}(p, \hat{p}) - M(p, \hat{p}) \right)$$

Showing $\pi^{(2)}(p, \hat{p})$ is positive is equivalent to showing that $\frac{\pi^{(2)}(p, \hat{p})}{M(p, \hat{p})}$ is positive since $M^{(2)}(p, \hat{p})$ is always positive.

Note that

$$\frac{\pi^{(2)}(p, \hat{p})}{M(p, \hat{p})} = (R(p) - w(\hat{p})) \frac{M^{(2)}(p, \hat{p})}{M(p, \hat{p})} - w'(\hat{p})$$

$$= (R(p) - R(\hat{p})) \frac{M^{(2)}(p, \hat{p})}{M(p, \hat{p})} + w'(\hat{p}) \left( \frac{M^{(2)}(p, \hat{p})}{M(p, \hat{p})} \frac{M(\hat{p}, \hat{p})}{M^{(2)}(\hat{p}, \hat{p})} - 1 \right)$$

At this point we use the assumption that $V(x) = P(x) = x$ for all $x$. Then after calculating integrals we obtain:

$$\frac{M^{(2)}(p, \hat{p})}{M(p, \hat{p})} = \frac{\int_{p}^{1} \frac{2\delta P'(p) V'(v)}{(\delta + P(v) - P(\hat{p}))^3} \, dv}{\int_{p}^{1} \frac{\delta P'(v)}{(\delta + P(v) - P(\hat{p}))^2} \, dv} = \frac{2\delta + 1 + p - 2\hat{p}}{(\delta + 1 - \hat{p})(\delta + p - \hat{p})} \cdot$$

Also,

$$\frac{M(\hat{p}, \hat{p})}{M^{(2)}(\hat{p}, \hat{p})} = \frac{\delta(\delta + 1 - \hat{p})}{(2\delta + 1 - \hat{p})}$$
Using these in the above implies

\[
\frac{\pi^{(2)}(p, \hat{p})}{M(p, \hat{p})} = \frac{(R(p) - R(\hat{p}))}{p - \hat{p}} \frac{2\delta + 1 + p - 2\hat{p}}{(\delta + 1 - \hat{p})(\delta + p - \hat{p})} - w'(\hat{p})(p - \hat{p}) \frac{\delta + 1 - \hat{p}}{(2\delta + 1 - \hat{p})(\delta + p - \hat{p})}
\]

\[
\geq \frac{1}{\delta + p - \hat{p}} \left( \frac{(R(p) - R(\hat{p}))}{\delta + 1 - \hat{p}} \frac{2\delta + 1 - \hat{p}}{\delta + p - \hat{p}} - w'(\hat{p})(p - \hat{p}) \frac{\delta + 1 - \hat{p}}{2\delta + 1 - \hat{p}} \right)
\]

since we assume \( p > \hat{p} \). Thus it is sufficient to show that for all \( p > \hat{p} \), that

\[
\left( \frac{(R(p) - R(\hat{p}))}{\delta + 1 - \hat{p}} \frac{2\delta + 1 - \hat{p}}{\delta + p - \hat{p}} - w'(\hat{p})(p - \hat{p}) \frac{\delta + 1 - \hat{p}}{2\delta + 1 - \hat{p}} \right)
\]

is positive.

Condition (2) implies that for all \( p \),

\[
w'(p) = (R(p) - w) \frac{2\delta + 1 - p}{\delta(\delta + 1 - p)}
\]

Substituting for \( w'(\hat{p}) \) into the above expression yields the following:

\[
(R(p) - R(\hat{p})) \frac{2\delta + 1 - \hat{p}}{\delta + 1 - \hat{p}} - (p - \hat{p}) \frac{R(\hat{p}) - R(\hat{p})}{\delta}
\]

Thus to show that \( \pi^{(2)}(p, \hat{p}) \) is positive, it is sufficient to show that

\[
\frac{R(p) - R(\hat{p})}{p - \hat{p}} - \frac{\delta + 1 - \hat{p}}{\delta} R(\hat{p}) - w(\hat{p}) \geq 0
\]

We concentrate on the case when firms with higher type have sufficient incentives to bid more than firms with lower types, i.e. a monotone equilibrium exists. From the analysis of the two-type case one suspects that this holds when the productivity of higher type firms are much higher than that of the lower types. In a continuous type space model this condition can be captured by assuming that the marginal productivity from increasing \( p \) is increasing, i.e. that \( R \) is a convex function. Under that assumption

\[
\frac{R(p) - R(\hat{p})}{p - \hat{p}} \geq R'(\hat{p})
\]

and thus it is sufficient to show that

\[
R'(\hat{p}) \geq \frac{\delta + 1 - \hat{p}}{2\delta + 1 - \hat{p}} \frac{R(\hat{p}) - w(\hat{p})}{\delta}
\]

To prove that this condition holds, first note that equation (2) implies that

\[
R(x) - w(x) \leq R'(x)\delta
\]
for all $x$. Then it follows that

$$R'(\hat{p}) > \frac{\delta + 1 - \hat{p}}{2\delta + 1 - \hat{p}} R'(\hat{p}) \geq \frac{\delta + 1 - \hat{p}}{2\delta + 1 - \hat{p}} \frac{R(\hat{p}) - w(\hat{p})}{\delta},$$

which shows that our condition is satisfied for any $\delta$.

### 6.2 CDF under uniform worker and firm types

Let $T_v(p) = Pr(wage \leq w(p) | type = v) = \frac{\delta}{\delta + v - p}$, if $v > p$ and 1 otherwise. The associated pdf is as follows:

$$t_v(p) = \frac{\delta}{(\delta + v - p)^2}$$

if $v < p$ and zero otherwise. Integrating over worker types the mass of firms of type $p$ that employ a worker is

$$m(p) = \int_p^1 \frac{\delta}{(\delta + v - p)^2} dv = \frac{1 - p}{1 - p + \delta}.$$  

To turn this into a probability density of firm type, we have to scale up to account for unemployed workers,

$$g(p) = \frac{1 - p}{1 - p + \delta} \cdot \frac{1}{1 - U}$$

where $U$ is the total mass of workers who are unemployed. Explicitly,

$$U = \int_0^1 \frac{\delta}{\delta + v} dv = \delta(\log(\delta + 1) - \log(\delta)) = \delta \log(1 + \frac{1}{\delta}).$$

This implies that the CDF associated with firm types in this economy looks like,

$$G(p) = \int_0^p g(x) dx = \int_0^p \frac{1 - x}{1 - x + \delta} \cdot \frac{1}{1 - U} dx$$

which can be solved as,

$$G(z) = \frac{z - \delta \log(\delta + 1) + \delta \log(-z + \delta + 1)}{1 - \delta \log(1 + \frac{1}{\delta})}. $$

---

*27To see this observe that

$$R(0) - w(0) = 0$$

and whenever

$$R(x) - w(x) = \delta R'(x)$$

holds it follows that

$$w'(x) = R'(x) \frac{2\delta + 1 - x}{\delta + 1 - x} > R'(x).$$

Then at that point $R - w$ is decreasing, while the right hand side, $\delta R'$ is increasing because $R$ is convex.
References


