



The University of Adelaide
School of Economics

Research Paper No. 2012-04
July 2012

Aggregate Reallocation Shocks and the Dynamics of Occupational Mobility and Wage Inequality

Jacob Wong



Aggregate Reallocation Shocks and the Dynamics of Occupational Mobility and Wage Inequality

Jacob Wong*

The University of Adelaide

This version : June, 2012

Abstract

This paper presents a dynamic model of structural unemployment and occupational choice in which an economy is subjected to aggregate reallocation shocks. Reallocation shocks, which change the relative labour productivity across occupations, drive variation in the distribution of workers across occupations. The wage paid to workers in a given occupation depends on its labour productivity and the number of workers employed in that occupation. Workers who wish to switch occupations in order to obtain higher wages face a fixed cost to retrain and, in addition, it is more costly to switch to occupations requiring vastly different skills relative to those of the worker's current occupation. Thus workers may prefer to remain unemployed in occupations suffering through relatively low productivity states. Between the late-1970s and the mid-2000s the U.S. economy featured an episode during which occupational mobility rose along with an increase in wage inequality both in the top and bottom halves of the wage distribution. This was followed by an episode during which occupational mobility fell, while a rise in inequality in the top half of the wage distribution was accompanied by a fall in inequality in the bottom half. The model can produce episodes with properties similar to that of the U.S. experience and thus offers a theory of why these episodes occur.

JEL CLASSIFICATION: E24, E32, J24, J31, J62

*School of Economics, The University of Adelaide, 3 & 4 Floor, 10 Pulteney St, Adelaide, SA 5005, Australia. Email : jacob.wong@adelaide.edu.au. Thanks to Paul Beaudry, Chris Edmond, Aspen Gorry, William Hawkins, Viktoria Hnatkovska, Andreas Hornstein, Nir Jaimovich, Amartya Lahiri, Henry Siu, and Mark Weder for valuable comments. Especial gratitude to Lawrence Uren for insightful remarks. I also thank presentation participants at the University of Adelaide, the University of Melbourne, UNSW, the Reserve Bank of Australia, the Federal Reserve Bank of Richmond, the Australian Conference on Quantitative Macroeconomics and the 2011 SED Meetings. Thanks to Thomas Lemieux, Giuseppe Moscarini and Chris Robinson for providing useful data. Finally, this paper grew out of collaborative work with Fabrice Collard to whom I am grateful.

It is now well known that in the U.S. economy there were large increases in both wage inequality and occupational mobility rates between the mid-1970s and the end of the 1980s. Interestingly, while occupational mobility rates dropped between the mid-1990s and the early-2000s, wage inequality in the top half of the wage distribution, as measured by the log 90/50 ratio (i.e. the ratio of the 90th percentile wage and median wage), continued to rise while wage inequality in the bottom half of the wage distribution, measured by the log 50/10 ratio, flattened or even decreased.¹ This polarization in wage inequality since the mid-1990s combined with the earlier dynamics of the wage distribution has resulted in divergent views concerning the causes of the changing dynamics of the wage structure ranging from changes in minimum wages to skill-biased technological change to offshoring of jobs. Recent work by Acemoglu and Autor (2010) and Autor, Katz, and Kearney (2008) suggest that the relative supply of workers to perform various tasks coupled with changing demands for task performance in the production process is important in determining both the distribution of wages across the aggregate economy and the allocation of workers across production tasks.

In this paper, I show that a simple dynamic general equilibrium model with costly occupational switching and occupational productivity shocks can be consistent with the stylized observations discussed above. Importantly, the model economy is only subjected to shocks which change the relative labour productivity across occupations. I refer to this type of shock as a reallocation shock. Furthermore, a measure of distance is introduced between occupations in order to capture the notion that switching between occupations that require the performance of vastly different tasks is more costly to workers. It turns out that the history of the distribution of workers across occupations plays a primary role in understanding the dynamics of wage inequality and mobility dynamics across both episodes. Given these simple assumptions, the model can generate episodes which are characterized by rising trends in the occupational mobility rate, as well as in wage inequality in the top and bottom halves of the wage distribution, as measured by the the log 90/50 ratio and log 50/10 ratio, respectively. The model can also generate episodes which are characterized by falling trends in the occupational mobility rate and wage inequality in the bottom half of the wage distribution along with a rising trend in wage inequality in the top half of the wage distribution. Combining specific assumptions about the initial distribution of workers across occupations along with the sequence of aggregate reallocation shocks hitting the economy, the model offers a theory to help understand the rise in both the occupational mobility rate and wage inequality from the mid-1970s through the mid-

¹The polarization in wage inequality since the mid-1990s is documented by Autor, Katz, and Kearney (2008) and Lemieux (2010). Moscarini and Thomsson (2007) and Robinson (2010) document the dynamics in occupational mobility and while the measured levels of mobility differs across their papers, the large rise and fall in trend occupational mobility is similar. Furthermore, their documentation of the rise in occupational mobility from the late-1970s through the late-1980s is consistent with the findings of Kambourov and Manovskii (2009a).

1980s as well as the subsequent fall in the occupational mobility rate which was accompanied by a polarization in wage inequality.

In the model, workers are distributed across a continuum of occupations with each occupation producing an intermediate good that may be used in the production of a final consumption good. At any point in time a given worker possesses skills to work in a particular occupation. Every occupation has its own level of labour productivity and all occupations experience fluctuations in labour productivity over time. Each period, a given worker faces an idiosyncratic fixed cost of retraining. Furthermore, if a worker chooses to retrain, the cost of switching to an occupation that uses different tasks relative to the worker's current occupation is high in comparison to switching to an occupation using similar tasks relative to the worker's current occupation. Workers who are skilled in an occupation where productivity is sufficiently high to pay a participation wage face the option of working or not working for the period and retraining in order to have the skills to work in a different occupation next period. Workers who are skilled in an occupation where productivity is not high enough to pay a participation wage face the option of being rest unemployed for the period and remaining in the occupation next period versus retraining and having skills to work in a different occupation next period. Over time, the economy is buffeted by reallocation shocks which may be interpreted as technological innovations that favour some occupations at the expense of others. The effect of the fixed costs of retraining as well as convex costs of retraining to dissimilar occupations is that the distribution of workers across occupations changes over time. This in turn affects the wage distribution across occupations. The distribution of wages feeds back into the equilibrium decisions of workers to retrain generating an interesting interplay between the distribution of employment across occupations, the wage distribution and the unemployment rate.

Solving for the equilibrium of this economy requires the ability to keep track of the distributions of workers across occupations as well as the distribution of labour productivity across occupations. I make two assumptions in order to manage the computational difficulties of tracking these two infinite dimensional objects. The first assumption is that occupations are located on the circumference of a circle; in other words, there is an occupation ring. The greater the distance between two occupations on the ring, the more dissimilar the tasks used by the occupations in production. The second modeling assumption is that the relative productivity across occupations in relation to the most productive occupation is preserved across time even though the identity of the most productive occupation can change over time. In this specific sense, a reallocation shock can be modeled simply as a change in the identity of the most productive occupation. Once the identity of the most productive occupation is known the relative productivity of all other occupations can be determined. These modeling assumptions reduce the computational problem to that of tracking the distribution of workers across occupations

and tracking the identity of the most productive occupation.

Using this model, I ask what conditions would have given rise to the qualitative behaviour in the trends of wage inequality and occupation mobility rates as observed in the U.S. between the mid- to late-1970s and the early-2000s. Specifically, within the confines of the model, I try to understand the properties of the distribution of workers across occupations at the beginning of this period along with properties of the sequence of aggregate reallocation shocks required to bring about the observed joint dynamics between wage inequality and occupational mobility rates. I argue that at the beginning of this period, a large mass of workers were attached to mid-level productivity occupations and so there were many middle-income earners. Then there was a large shift in the relative productivity across occupations such that relative productivity in occupations which were mid-level productivity occupations, dropped continuously for almost a decade. As workers chased higher wages by switching towards more productive occupations, this gave rise to an increase in occupational mobility rates. However, on account of occupation switching frictions (say fixed costs of retraining, etc.) not all individuals switched occupations immediately. Many workers remained attached to their occupations and these workers either experienced large drops in their wages or became unemployed. This “stickiness” in the distribution of workers across occupations caused an increase in the mass of workers at low productivity occupations and dragged down the 10th percentile wage faster than the median wage (which also dropped during this period) so that wage inequality in the bottom half of the wage distribution increased. Simultaneously, there was a small set of workers who were attached to occupations which experienced relative productivity gains and these workers saw increases in their wages. Therefore, given the drop in the median wage, there was an increase in wage inequality in the top half of the wage distribution.

In order to explain the observed polarization of wage inequality and accompanying decrease in occupational mobility rates between the mid-nineties and the early-2000s, the model requires that there were no drastic changes in relative productivity across occupations over this decade. By the end of the first episode, workers were dispersed across a much larger set of occupations in comparison to the distribution in the late-1970s. Many workers were attached to occupations in which productivity was very low relative to the most productive occupation. Over time, workers who were unemployed or attached to low-paying occupations acquired skills to work in different occupations. Thus occupation mobility rates were high at the beginning of this decade and began waning as more and more workers retrained and then remained attached to their new occupations. This resorting of workers across occupations in the absence of significant reallocation shocks caused the downward trend in occupational mobility rates over this period. Furthermore, as workers at the bottom end of the productivity spectrum reallocated themselves to higher productivity occupations, individually they became better off as their wages rose.

However, as many of these workers entered similar occupations they put downward pressure on the wages in this set of occupations simply by increasing the supply of workers in these occupations. Hence as the wage earned by the workers in the lower percentiles of the wage distribution increased, the bunching of workers into mid-level productivity occupations caused the wage earned by the median wage earner to remain stable relative to rising wages at the bottom of the wage distribution. This was responsible for the drop in wage inequality between the median wage earner and the 10th percentile wage earner. As at the bottom end, workers at the top end of the distribution also switched into more productive occupations. Given a small mass of workers in these high-productivity occupations the downward pressure on wages emanating from the increase in supply of workers in this subset of high paying occupations did not dominate the gains from increased productivity and upward pressure on wages from higher aggregate output. This increase in the wages at the top of the distribution also caused a rise in wage inequality between the 90th percentile wage earner and the median wage earner. Thus in the model, a polarization in wage inequality arises without any significant aggregate reallocation shocks.

The issues addressed in this paper overlap with that of a growing body of literature which focuses on the interaction of skills, tasks and technologies. This literature, as summarized by Acemoglu and Autor (2010), has tried to understand how technological innovations and task specific demand shocks shaped the dynamics of the wage distribution in the U.S. economy between the late 1970s and the mid-2000s. Arguments, notably by Autor, Levy, and Murnane (2003), have been put forth that technological innovations replaced routine tasks resulting in a decrease in demand for workers who perform such tasks. This hollowed out the middle of the wage distribution resulting in polarization. After extracting characteristics across a large set of occupations, Firpo, Fortin, and Lemieux (2010) argue that occupations using tasks most directly affected by technological change and offshorability drive much of the change in wage inequality between the 1980s and the 2000s. Acemoglu and Autor (2010) and Costinot and Vogel (2010) provide a model in which workers possess particular skills that may be used to perform various tasks in the production process. They use their model to study the effects of task-replacing technologies, and offshoring on the equilibrium wage distribution and supply of tasks. The model in this paper differs in that it focuses on technological innovations which affects relative labour productivity across occupations and the speed with which workers reallocate themselves across occupations. In focusing on these dynamic issues, the model purposely abstracts from features which revolve around a hierarchy of skills. Recent empirical work by Poletaev and Robinson (2008), Gathmann and Schonberg (2010) and Robinson (2010) construct measures of task-similarity across occupations and argue that workers who switch occupations tend to switch to occupations using similar tasks as their previous occupations. In light of these observations,

I impose a notion of relative distance between occupations but do not rank occupations in any sense which requires some occupations or tasks to require more “skill” than others. To my knowledge, this is the first equilibrium model in which there are aggregate fluctuations in the distribution of workers across a large number of occupations in which individuals choose the “distance” between current and past occupations. Thus this paper can be viewed as an examination of the joint dynamics of the occupational mobility rate and the wage distribution which works alongside contributions of previous work in this area.

From the point of view of studying the joint dynamics between occupational mobility and wage inequality, my work is most related to Kambourov and Manovskii (2009a). In addition to presenting some striking empirical observations on the rate of occupation switching its relationship with wage inequality, their paper presents a structural model calibrated to understand the change in wage inequality and occupational mobility between the period 1970-1973 and 1993-1996. An important feature of their model is that they tie the relationship between wage inequality and occupational mobility to a mechanism that generates wage increases due to occupational tenure. Kambourov and Manovskii (2009b) shows that occupational tenure is more important than job tenure as a determinant of wage growth. While extremely similar in spirit, this paper focuses on presenting the dynamics of an aggregate economy constantly being buffeted by aggregate shocks and so I concentrate on inter-occupational wage differences and omit the intra-occupational wage mechanism featured in Kambourov and Manovskii (2009a).

From the modeling perspective, the literature most related to the model in this paper is that which has grown from the seminal paper of equilibrium unemployment by Lucas and Prescott (1974). In their paper, Lucas and Prescott introduce a theory of equilibrium unemployment in which workers are attached to islands which, in my context, may be thought of as occupations. Occupations are subject to idiosyncratic demand shocks which affect the wages paid to workers attached to that occupation. Workers may choose to work at the prevailing wage, leave the occupation and search for new jobs or stay attached to occupations and wait for better times. The main difference with my model is that in my model there is a measure of distance between occupations and it is more costly for workers to move to occupations that use dissimilar tasks than the worker’s current occupation. In the classic Lucas-Prescott island model, when workers decide to switch islands, they sit out of the labour force for a period and then have unrestricted choice of islands upon which to reenter the labour force. Alvarez and Shimer (2009a) and Alvarez and Shimer (2009b) present extensions of the Lucas-Prescott model and using continuous-time dynamic programming techniques are able to present elegant closed-form solutions to a version of the economy’s steady state.² While both models provide insights into the theory of rest

²It is worth noting that Alvarez and Shimer (2009a) use their model to understand wage behaviour at the industry level whereas my application is to highlight the dynamics of the correlations between the unemployment rate, occupational mobility rates and measures of wage inequality.

unemployment, and solve for the steady state in which workers direct their choice of occupation switches, neither paper examines the dynamics of the economy when subjected to aggregate shocks. However, this paper shares the spirit of the literature built on the Lucas-Prescott island model.³

The idea of modelling distance using the circumference of a circle has been exploited in the equilibrium unemployment literature by Marimon and Zilibotti (1999) and Gautier, Tuelings, and Vuuren (2010). These papers use distance around a circle to characterize mismatch between a worker's skills and a firm's technology. In their models a worker is located at a point on a circle and is matched with a firm that is also located at a point on the circle. The greater the distance between the worker and the firm, the larger the mismatch and the lower is output generated by the worker-firm production unit. Clearly such a concept of mismatch carries a similar spirit as in this paper where the distance is used to characterize mismatch between a worker's current occupation specific skills and skills required for work in other occupations.

I now lay out a map for the rest of the paper. In Section 1 I point out some features in the U.S. data and then in Section 2 I describe a model that is built to help understand these observations. Section 3 presents the main quantitative properties of the model. Specifically, in order to illustrate the mechanisms at work in the model, the model's implications concerning the joint dynamics the occupational mobility rate and wage inequality are examined. A discussion concerning other uses of the model is provided in Section 4 and finally concluding remarks are offered in Section 5.

1 Observations from U.S. Economy

In this section of the paper I reproduce figures that illustrate some properties of the data noted in earlier work concerning occupational mobility rates and wage inequality from the late-1970s and the mid-2000s. In doing so, I highlight the joint dynamics of occupational mobility rates and measures of wage inequality over the mid-1990s and the early-2000s period which has received less attention to date. A well known feature of occupational mobility rates is that there was an upward surge between the late-1970s and the mid-1980s. The upper-left panel in Figure 1 displays monthly observations on occupational mobility as constructed by

³There has been recent work using the Lucas-Prescott island model in order to examine the effects of aggregate shocks on unemployment. Particularly, Gouge and King (1997) as well as Garin, Pries, and Sims (2011) use a two-island model to examine the relationships between productivity and unemployment. For the stylized observations that are examined in this paper, a richer set-up is required because in a two island model, the 50th percentile wage earner would necessarily earn the same wage as either the 90th percentile worker or the 10th percentile worker, so long as all workers on an island earned the same wage. Veracierto (2008) embeds an island model of employment into an otherwise standard real business cycle model and examines the effects of aggregate productivity shocks on unemployment rates and labour force participation rates. Carrillo-Tudela and Visschers (2011) use a many-island model to examine the effects of allowing frictional unemployment on each island and aggregate shocks on unemployment and occupational mobility rates.

Moscarini and Thomsson (2007). At the three-digit coding level for occupations as tabulated in the Current Population Survey (CPS), there is a notable rise in the occupation switching rate but this is followed by a sizeable decrease from the mid- to late-1990s through the early-2000s. The smooth line is the trend as extracted by a band pass filter leaving behind only frequencies longer than eight years.⁴ The argument for the eight year cut-off is that work on medium-run cycles often look at cycles between eight and fifty years in duration. As I concentrate on the medium-run, eight years seems like a reasonable cut-off frequency.

The upper-right panel of Figure 1 displays calculations from Robinson (2010) in which annual occupation switching rates are estimated at the 3-digit occupational coding level, as well as for switches that satisfy certain requirements for the change in a worker's skill portfolios. There is much anecdotal evidence that many occupation switches do not involve any change in tasks but rather are simple re-categorizations of workers' coded occupations. Poletaev and Robinson (2008) construct measures of distance between occupations by taking measures of a large number of characteristics by occupation, constructing a four-dimensional skill vector from these occupation characteristics and then imposing a measure of distance between these skills. The plot displays annual occupational switching rates at the 3-digit level, as well as measures which require that an occupation switch involve a change in the main skill used. The "Main Skill" employed by an occupation is the entry in the occupation's skill vector with the highest numerical value. A worker switches occupations by this measure if, in the four-dimensional skill vector, the element with the highest skill score changes with an occupation. The "PC1" measure counts an occupation switch if the main skill switches between occupations, the change in scores attached to the main skill in the pre-switch occupation is not too small and the level of this main skill in the pre-switch occupation is not too low.

In the model that is described later in the paper, workers must pay a fixed cost of switching occupations. If these fixed costs are substantial then workers will not tend to make trivial occupation switches in the sense of moving into occupations that are virtually identical in their task or skill demands. This is because, in the model, such occupations would pay near identical wages. Thus it is useful to present measures of occupation switches which account for some measure of distance between switches. Qualitatively, the pattern of occupation switching displayed in the annual measures present a similar pattern as those measured at the monthly frequency and clearly exhibits a downward trend from the mid-1990s onwards.

Previous work by Kambourov and Manovskii (2009a) noted that the large rise in occupation switching rates during the late-1970s through the 1980s was accompanied by a rise in overall wage inequality. Separately, empirical work such as those by Autor, Katz, and Kearney (2008)

⁴I smooth the data without gaps just to give a feel for the slow moving component of the data. A full description on the issues with comparing periods before and after the gaps is given by Moscarini and Thomsson (2007).

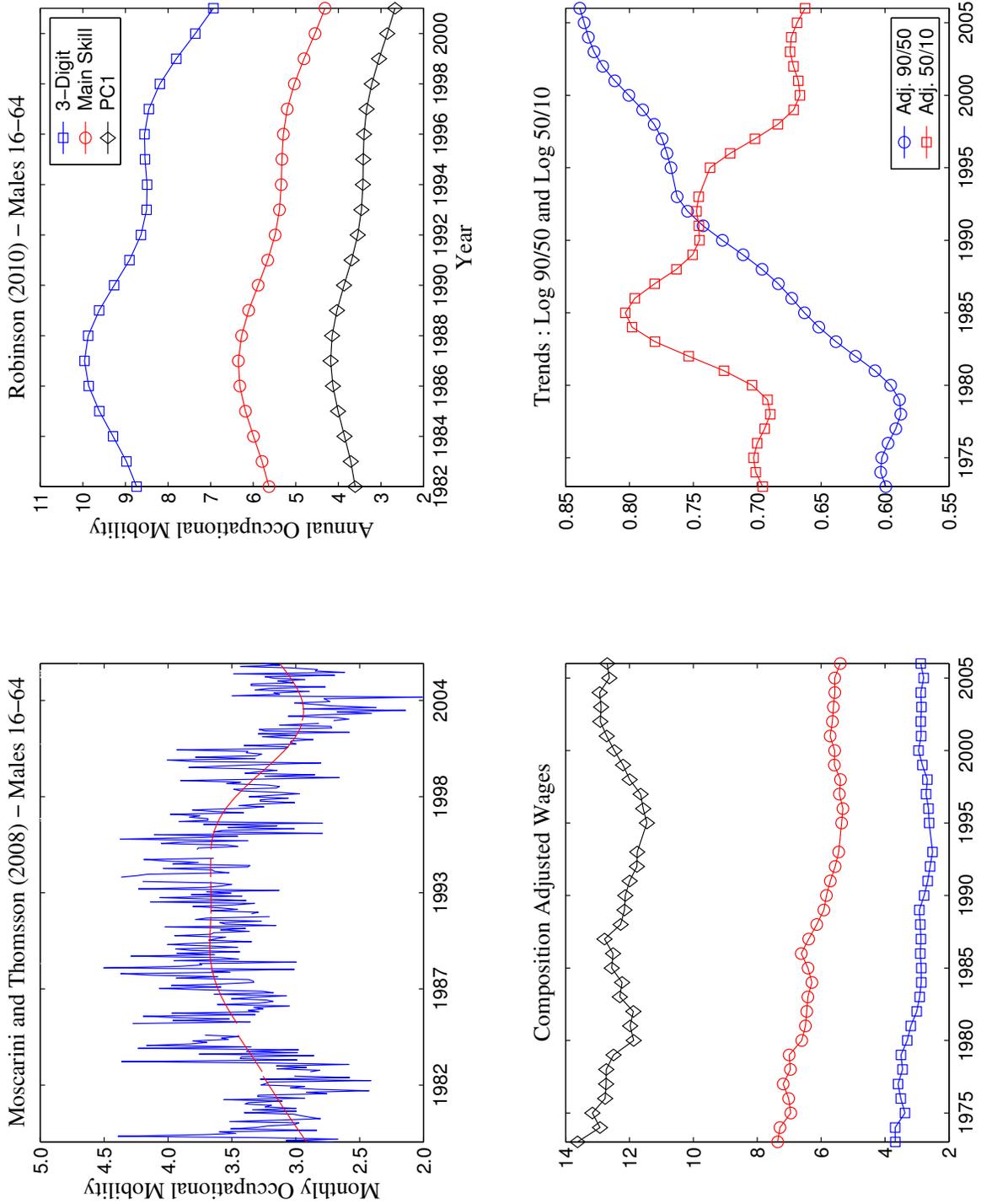


Figure 1: U.S. Economy - Occupational Mobility and Wage Inequality

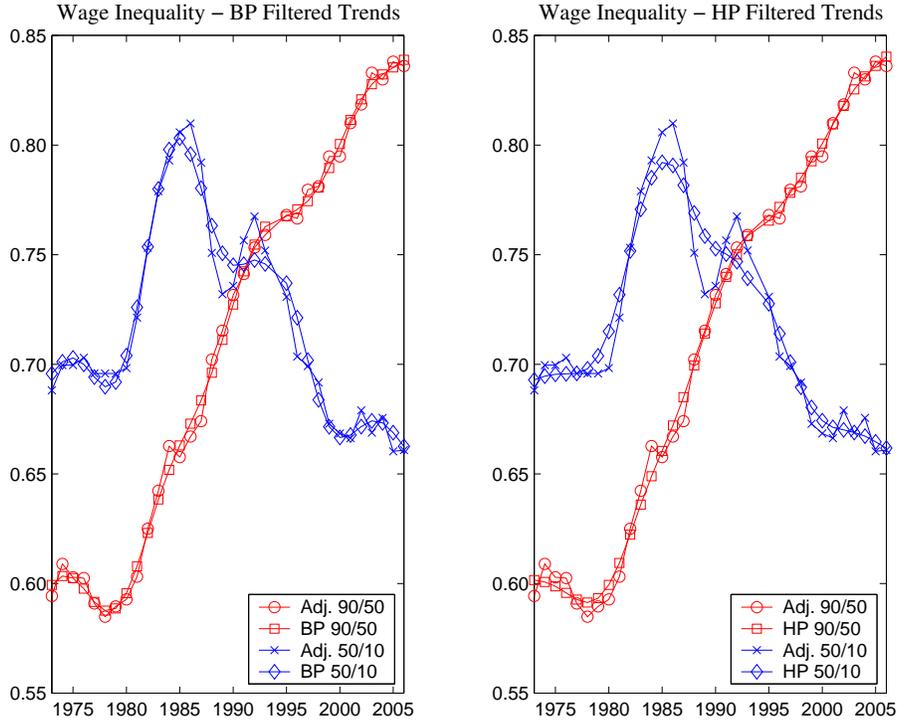


Figure 2: Wage Inequality - Obtaining the Trends

and Lemieux (2010) pointed out that while there was a rise in wage inequality in both the top half and bottom half of the wage distribution between the late 1970s and mid-1980s, there was clearly a change in this relationship after the mid-1980s as wage inequality in the top half continued rising while falling in the bottom half of the distribution. The bottom-right panel in Figure 1 shows annual measures of wage inequality as constructed by Lemieux (2010) after using a band pass filter to remove frequencies below eight years. The composition adjusted series displayed attempts to control for the distribution of work experience and education over time. It is commonly believed that accruing experience augments wages as does higher levels of education so controlling for these factors across the working population allows for a better understanding of the forces driving wage inequality. In my model, workers do not accumulate occupational specific capital, nor are workers distinguished by educational attainment so the composition adjusted data seems to be an appropriate reference. While the log 90/50 ratio grows consistently from 1979 to 2006 the log 50/10 ratio peaks in 1985 and then, with a minor blip in the early 1990s, falls through to 2006.

Figure 2 contrasts the trends extracted using the band pass filter at medium run frequencies to those extracted by the HP filter with smoothing parameter 6.25, which is frequently used when applying the filter to annual data. Both panels plot the extracted trends against the raw

data. Notice that the blips in the trend which occur in the early 1990s and the early 2000s through use of the band pass filter are not present when using the HP filter. It can be argued when eye-balling the raw data that the downward trend in the log 50/10 ratio begins in 1985 and continues uninterrupted until 2001 whereas the uninterrupted downward trend in the band pass filtered data begins in 1992 and ends in 2000. There is also an uninterrupted downward trend from 1985 through to 1990 when looking at the band pass filtered data. In contrast examination of the HP filtered trend for the log 50/10 ratio reveals an uninterrupted downward trend from 1985 through 2006. While the length of the uninterrupted downward trend may be debated as ranging from 7 to 8 years or 21 years, there is no argument that there were extended periods of uninterrupted decreases in wage inequality in the bottom half of the wage distribution.

In order to gain a further understanding of the behaviour of the wage distribution, the bottom-left panel of Figure 1 plots the composition adjusted log wage for the 90th, 50th and 10th percentile wage earners. It can be seen that during the episode when the trends in the log 90/50 and the log 50/10 ratios both rose, wages at the top rose while the median wage fell and wages at the bottom fell faster than the median wage fell. In contrast during the wage polarization episode of the mid-1990s to the early 2000s, wages at the top rose faster than the median wage and wages at the bottom also rose faster than did the median wage.

Summarizing, there was a rise in wage inequality in both the top and bottom halves of the wage distribution from 1979 through to the mid-1980s and this was accompanied by an increase in the occupational switching rate. From the early-1990s through to the early-2000s there was a continued increase in wage inequality in the top half of the wage distribution accompanied by decreases in wage inequality in the bottom half of the wage distribution and in the rate of occupational switching. I now present a model that is used to help shed some light on the observed dynamics in the data.

2 The Model

I construct a model in which technological innovations constantly shifts relative labour productivity across occupations. Workers must choose whether to incur the costs of retraining to work in new occupations, to be unemployed if no work is available in his/her current occupation or to work at the prevailing wage in his/her current occupation. For simplicity, workers only possess the skills to work in one occupation at any point in time. The timing is as follows: at the beginning of the period, workers are distributed across a continuum of occupations. Each worker draws an idiosyncratic fixed cost of switching occupations in the current period. The distribution of labour productivity across occupations in the current period is then revealed and then workers decide whether to switch occupations, or stay in their current occupation.

Production, trade and consumption occurs and the period ends. In the following subsections, details regarding the economy’s production structure, the problem of the individual worker and the definition of an equilibrium are provided.

2.1 Technology

There is a continuum of occupations on a circle with radius one indexed by $i \in [0, 2\pi]$.⁵ Consider a uni-modal function $g(\cdot)$ that is symmetric around the mode with domain $[-\pi, \pi]$.⁶ Let θ denote the location of the mode on the circumference in the current period and assume that θ follows the process⁷

$$\theta' = \begin{cases} \theta + \epsilon' & \text{if } \theta + \epsilon' \in [0, 2\pi] \\ \theta + \epsilon' - 2\pi & \text{if } \theta + \epsilon' > 2\pi \\ \theta + \epsilon' + 2\pi & \text{if } \theta + \epsilon' < 0 \end{cases}, \quad \epsilon \sim F[-\pi, \pi].$$

Let $F(\epsilon)$ be continuously differentiable on $(-\pi, \pi)$ and denote its density by $f(\epsilon)$. Let the relative location of $i \in [0, 2\pi]$ from the location of the mode at time t be given by

$$\delta(i) = \min\{\theta - i, 2\pi - \theta + i\}$$

such that the distance of i from θ is $|\delta(i)| = \min\{|\theta - i|, |2\pi - |\theta - i||\}$. Let the value $A(i) = g(\delta(i))$ be the level of labour productivity in occupation i given that occupation i has a location $\delta(i)$ relative to the location of the productivity frontier. I assume that $g(\cdot)$ is twice continuously differentiable on $[-\pi, \pi]$. Denote the height of $g(\cdot)$ at the mode by \bar{A} so by construction, $g(0) = \bar{A}$. Finally, assume that $g(\cdot) \in [0, \bar{A}]$ with $\lim_{\delta \rightarrow -\pi} g(\delta) = \lim_{\delta \rightarrow \pi} g(\delta) \cong 0$.

On the left side of Figure 3 shows an example in which the mode shifts from one period to the next while the function $g(\delta)$ retains its shape. The height of the function $g(\cdot)$ at any point on the circumference of the occupation ring denotes the labour productivity of the occupation located at that point. As reallocation shocks swing the mode of the $g(\cdot)$ function around the ring, the labour productivity of any occupation also changes but the productivity of occupations a given distance from the productivity frontier (i.e. the location of the mode of $g(\cdot)$) remains constant over time. The right side of Figure 3 depicts an example of the determination of $\delta(i)$ for occupation i when θ shifts over time. Given the assumption that $g(\cdot)$ is uni-modal, the only relevant property to determine an occupation’s labour productivity is its current distance from the productivity frontier.

⁵One example would be that each occupation requires a particular set of skills of its workers $\{x, y\}$ with $x \in [-1, 1]$ and $y \in [-1, 1]$ in order to perform an occupational-specific task. All existing occupations are such that $\sqrt{x^2 + y^2} = 1$. Then to work in a given occupation workers must possess the skill bundle demanded to perform the task in that occupation.

⁶The assumption of a uni-modal function is made for simplicity and is not essential. As discussed later, the essential property is that the function $g(\cdot)$ possesses a tractable reference point.

⁷I follow the convention that “primed” variables indicate the value of the variable in the next period.

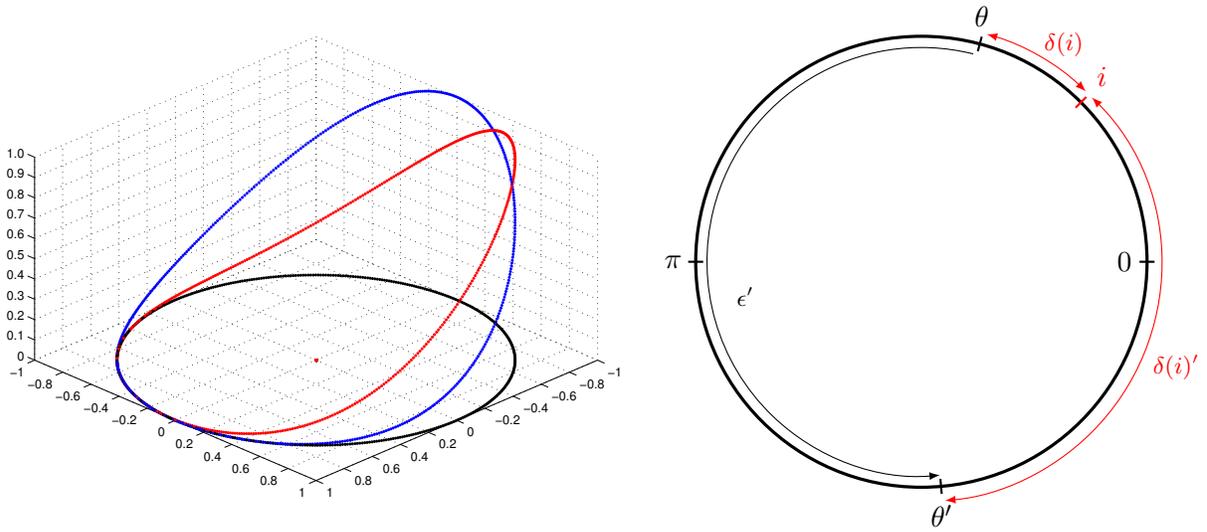


Figure 3: Reallocation Shocks and Determining the Value of $\delta(i)$

The assumption that the function $g(\delta)$ is time-invariant can be relaxed but for this paper, the assumption is held in order to understand the qualitative properties of model economy as a first pass. One way to think about a reallocation shock is that there is a technological innovation that favours a particular subset of occupations (or tasks) relative to other occupations. For example, computer innovations have been argued by Autor, Levy, and Murnane (2003) to have favoured occupations requiring non-routine tasks as opposed to occupations requiring routine tasks, say many of the occupations in the manufacturing sector. As these innovations favoured particular types of tasks, all occupations using similar tasks would gain in productivity, though some more than others.

2.2 Production

2.2.1 Final Good Firms

There is an aggregate consumption good which is produced by perfectly competitive final good firms. The final good firms use the output of intermediate good firms as inputs. Intermediate good i can only be produced by firms in occupation i where $i \in [0, 2\pi]$. Final good firms take the price of intermediate goods from occupation i as given and the price of the final good is normalized to one. Notice that given the assumptions on productivity, the output of a firm-worker pair is independent of the occupation's identity; it is only dependent on the distance of the occupation from the most productive occupation. Thus, for an occupation that is δ from the most productive occupation, I can write the output of a worker-firm pair as $A(\delta)$. Denote the measure of workers attached to an occupation δ from the productivity frontier and the measure

of employed workers in that occupation in the current period by $\psi(\delta)$ and $n(\delta, \psi)$ respectively. If there is production from the occupation δ from the frontier in the current period, then total output from this occupation is denoted by $y(\delta, \psi) = A(\delta)n(\delta, \psi)$.

Final good firms aggregate across all available intermediate goods using the production function

$$y(\psi) = \left[\int_{\Omega} y(\delta, \psi)^{\frac{\chi-1}{\chi}} d\delta \right]^{\frac{\chi}{\chi-1}}$$

where Ω is the set of all occupations with strictly positive output in the current period, and $\chi > 0$ is the elasticity of substitution between intermediate goods.

2.2.2 Intermediate Good Firms

Intermediate good i is produced by worker-firm pairs in occupation i through use of a linear production technology, such that, in the current period, each employed worker in occupation i produces an amount $A(i)$. Again, given the modeling assumptions I can rewrite this in terms of intermediate goods δ from the frontier so that $\delta \in [-\pi, \pi]$ and output of a worker-firm pair in an occupation δ from the frontier in the current period is $A(\delta)$. There are no set-up costs of period production which means that as long as production is profitable, firms continue to enter any given occupation until all workers in that occupation are hired. Suppose that workers have an outside option relative to employment of b if they stay at home. This period payoff to unemployment, b , can be either be thought of as output from home production or as unemployment benefits. The minimum period operating costs to a firm is b as firms have to cover the participation wage of their workers.

Let the price of intermediate good produced by the occupation δ from the frontier, given the distribution of workers $\psi(\delta)$ be $p(\delta, \psi)$. All intermediate good firms take the price of their output as given. The inverse demand function for any firm operating in an occupation δ from the frontier is given by

$$p(\delta, \psi) = \left[\frac{y(\delta, \psi)}{y(\psi)} \right]^{\frac{-1}{\chi}} \quad \forall \delta \in \Omega,$$

where I have normalized the price of the final good to be one.

2.3 Worker's Problem

There are no savings in this economy and a worker's period utility from consumption is given by $u(c)$ with $u' > 0$, $u'' < 0$ and $u(0) > -\infty$. A worker who is located in occupation i may choose to be employed, if there are firms in occupation i , unemployed but attached to occupation i , or not participating in the labour force. Period consumption in the unemployed

state is $b(\psi) > 0$. Unemployment consumption is a function of the distribution of workers across occupations which allows for unemployment consumption to be a fixed fraction of the average wage in the economy. This case allows us to approach a notion of unemployment benefits as a replacement ratio. When a worker is not in the labour force, the worker may choose to move to a different occupation. When choosing to switch occupations a worker enjoys consumption in the amount $b(\psi)$ but must incur an idiosyncratic fixed utility cost of switching, z . Each period a worker's idiosyncratic cost of moving is drawn from a time invariant distribution $H(z)$ with density $h(z)$ and support $[0, \infty)$. Assume that $H(z)$ is twice continuously differentiable everywhere and that $\int_0^\infty z dH(z)$ is bounded from above. Furthermore, there is also a convex disutility cost of retraining, $\varphi(x) \geq 0$, where x is the measure of distance between the worker's current occupation and his/her desired, new occupation. Simply, x is the distance around the circumference of the occupation ring between the old and the new occupation. Assume that $\varphi(0) = 0$ so the cost of moving zero distance is zero. Each worker faces the probability of death, $1 - \beta$, at the end of each period. In order to keep the size of the population at one, a measure $1 - \beta$ of new workers are born at the end of each period. New born workers are uniformly distributed across all occupations.

When a worker is employed in an occupation that is a distance δ from the frontier, the period wage is $w(\delta, \psi)$ which yields period utility $u(w(\delta, \psi))$. Let the value of being employed in an occupation that is δ units away from the productivity frontier be $V(\delta, \psi)$. Denote the value of being unemployed in an occupation δ from the frontier to be $U(\delta, \psi)$ and denote the value of being a non-participant or occupation switcher (and retraining) to be $T(\delta, \psi)$. In this model there is no "search unemployment" as all workers who are unemployed are not technically looking for jobs because they know that there are no jobs available in their occupation. In some sense, all workers who are unemployed have been displaced and are not looking for a job so in the spirit of Alvarez and Shimer (2009a), I refer to these workers as "rest unemployed". Due to the nature of unemployment in my model, this paper can be viewed as complementary to papers in which unemployment arises due to search frictions.⁸ I refer to workers in the retraining state as non-participants because they are not actively seeking a job.

It is clear that as competition amongst firms results in firms paying at least the participation wage, workers will never choose to quit a job to enter unemployment; if workers quit then they must be choosing to retrain. Therefore the only time a worker enters the unemployment pool is if the worker is displaced, that is, firms in the given occupation shut down and the worker

⁸For example, papers such as Burdett, Shi, and Wright (2001), Peters (2000) and Shimer (2005) provide a theory of unemployment due to coordination frictions between workers and firms with job vacancies. Models of competitive search as in Moen (1997) and the popular random matching models described in Pissarides (2001) are less specific regarding the nature of the matching frictions in the labour market and may capture the structural unemployment that I model explicitly in this paper.

does not choose to retrain. However, workers may choose to quit in order to retrain.

Let $\Omega := \{\delta : y(\delta) > 0\}$, so that Ω is the set of all occupations in which production is strictly positive. Then

$$\begin{aligned} V(\delta, \psi) &= u(w(\delta, \psi)) + \beta \left\{ \int_{[-\pi, \pi] \setminus \Omega'} \int_0^\infty \max \{T(\delta', \psi') - z', U(\delta', \psi')\} dH(z') dF(\delta'|\delta) \right. \\ &\quad \left. + \int_{\Omega'} \int_0^\infty \max \{V(\delta', \psi'), T(\delta', \psi') - z', U(\delta', \psi')\} dH(z') dF(\delta'|\delta) \right\}. \end{aligned} \quad (1)$$

Workers who are unemployed are faced with the decision to stay unemployed in their current occupation or to exit the labour force and retrain. The value function for an unemployed worker in a occupation δ away from the frontier is

$$\begin{aligned} U(\delta, \psi) &= u(b(\psi)) + \beta \left\{ \int_{[-\pi, \pi] \setminus \Omega'} \int_0^\infty \max \{T(\delta', \psi') - z', U(\delta', \psi')\} dH(z') dF(\delta'|\delta) \right. \\ &\quad \left. + \int_{\Omega'} \int_0^\infty \max \{V(\delta', \psi'), T(\delta', \psi') - z', U(\delta', \psi')\} dH(z') dF(\delta'|\delta) \right\}. \end{aligned} \quad (2)$$

Finally, workers who are non-participants will incur a fixed cost to retrain. The retraining decision comes at a utility cost of z in addition to a convex utility cost, $\varphi(x)$, in the distance traveled, x , which denotes the measure of occupations that the worker passes over in the period. Thus, for non-participants, $\delta' = \delta + x + \epsilon'$ allowing the value function for a non-participant who is currently δ from the productivity frontier to be written as

$$\begin{aligned} T(\delta, \psi) &= \max_x \{u(b(\psi)) - \varphi(x) \\ &\quad + \beta \left\{ \int_{[-\pi, \pi] \setminus \Omega'} \int_0^\infty \max \{T(\delta', \psi') - z', U(\delta', \psi')\} dH(z') dF(\delta'|\delta + x) \right. \\ &\quad \left. + \int_{\Omega'} \int_0^\infty \max \{V(\delta', \psi'), T(\delta', \psi') - z', U(\delta', \psi')\} dH(z') dF(\delta'|\delta + x) \right\}. \end{aligned} \quad (3)$$

Note that $T(\cdot)$ is the value of switching occupations after having incurred the fixed cost of switching occupations. Conditional on choosing to switch occupation, a retraining worker chooses to move along the occupational ring if the marginal cost of moving an extra unit of distance is less than the marginal expected increase in value of being attached to an occupation a unit further along the occupational ring.

2.4 Equilibrium

Definition 1 *An equilibrium is*

1. a set of value functions $V(\delta, \psi)$, $U(\delta, \psi)$, and $T(\delta, \psi)$, along with the decision rule for relocation, $x(\delta, \psi)$, and choice of employment states that satisfy the Bellman equations (1), (2), and (3),
2. a distribution function $\psi(\delta)$ consistent with the individual's decision rules and satisfying a transition function $\psi' = \Gamma(\psi, \epsilon')$,
3. a price function $p(\delta, \psi)$ that clears all markets in which there is output,
4. an unemployment consumption function $b(\psi)$ which is consistent with optimal decisions of workers, and
5. where there is production, a wage function $w(\delta, \psi)$ such that firms always obtain zero profits.

Free entry into each occupation causes the worker's wage to be equal to total revenues of the firm, $p(\delta, \psi)A(\delta)$, so that $w(\delta, \psi) = p(\delta, \psi)A(\delta)$. Thus firms in occupation $\delta \in \Omega$ pay their workers a wage of

$$w(\delta, \psi) = A(\delta)^{\frac{\chi-1}{\chi}} n(\delta, \psi)^{-\frac{1}{\chi}} y(\psi)^{\frac{1}{\chi}}.$$

It need not be the case that $n(\delta, \psi) = \psi(\delta)$ as some people who are attached to occupation δ may choose to switch occupations. Notice that pushing χ to infinity, the wage becomes equal to labour productivity only; the measure of workers in an occupation no longer matters for the wage. I will refer to this case as the partial equilibrium model. For the discussion that follows, note that the wage ratio between any two occupations is positively related to the ratio of labour productivities between the two occupations and negatively on the ratio employment between the two occupations. In partial equilibrium, the wage ratio between two occupations is only a function of the relative productivities between the occupations.

3 Quantitative Results

I now examine whether the model can generate the type of trends that appear in the data. After a discussion regarding the parameterization of the model, an example is provided which is designed to illustrate how the model generates the type of dynamics observed in the U.S. economy between the mid-1970s and the early 2000s. Following the example, some simulation results are provided in order to highlight features of the distribution of workers across occupations and the sequences of shocks that are characteristic of different types of episodes which conform with those observed in the data. Finally, I check to see if the model can generate the two different types of episodes sequentially, as observed in the data, when the

Table 1: Parameter Values

| Parameter | Value | Description |
|-------------------|---------|--|
| β | 0.99375 | Subjective discount factor |
| σ | 1 | Coefficient of relative risk aversion |
| χ | 2.25 | Elasticity of substitution in final goods CES aggregator |
| \underline{a} | 0.01 | Lower bound for labour productivity |
| \bar{a} | 1 | Upper bound for labour productivity |
| ξ | 2.65 | Shape parameter for labour productivity function $g(\delta)$ |
| α_ϵ | 100 | Parameter governing distribution of shocks |
| μ_z | 4.00 | Mean of idiosyncratic shocks from Gamma Distribution |
| σ_z^2 | 2.50 | Variance of idiosyncratic shocks from Gamma Distribution |
| b | 0.33 | Fraction of average period wage enjoyed by unemployed |
| η | 60 | Weight on quadratic retraining cost |

aforementioned features of the distribution of workers and sequences of shocks are fed into the model.

3.1 Parameterization

The model is highly stylized and parsimonious in the number of parameters to be calibrated for the following reason. Several features of the model do not have counterparts in much of the previous literature so there are no existing values for some parameters which can be adopted from other work.⁹ In choosing the model's parameter values for the following computational exercise, I stray from the standard method of matching steady state targets. The reason for doing so is that the steady state is close to the lower bound of feasible unemployment rates and occupational switching rates in this economy. Intuitively, if the economy starts off in a stochastic steady state (i.e. a distribution of workers across occupations when the economy is continually hit with zero shocks) then any reallocation shock will result in an increase in both the unemployment rate and the occupational switching rate. Given this obstacle, I choose parameter values such that specific targets lie in the 95% band from an ensemble of simulations.¹⁰ I target particular statistics observed in U.S. data described below.

Table 1 displays the parameter values used in the numerical simulations. In the simulations that follow, a period represents three months so that the value of β is such that a worker is expected to have a work life of forty years. I choose utility from consumption to be given by

⁹The computational procedure used to approximate the model is provided in the Appendix.

¹⁰For a given set of parameters I solve the nonlinear model and simulate the economy 1000 times with 249 periods per simulation after a burn-in run of 500 periods is discarded. Then, for a particular statistic, I sort the ensemble of 1000 points and discard the lowest and highest 2.5%. I then check to see if the statistic from the U.S. data lies in this interval from the simulated ensemble.

$u(c) = \frac{c^{1-\sigma}-1}{1-\sigma}$ and fix $\sigma = 1$ so that the utility function is logarithmic. The maximum and minimum values of labour productivity are normalized to be $\bar{a} = 1$ and $\underline{a} = 0.01$. As long as $b > \underline{a}$ the lower bound of labour productivity plays no role in the analysis. As the final goods sector aggregates across intermediate goods produced by different occupations, there is little guidance on the choice of χ . Alvarez and Shimer (2009a) argue that in a model of sectoral mobility elasticities in the range of 2.8 to 4.5 may be reasonable. My model is not directly comparable to theirs as I consider occupational mobility and they consider sectoral mobility. Kambourov and Manovskii (2009a) consider a model in which all occupations produce a homogeneous good so the analogy in the model of this paper would be that intermediate goods would be perfectly substitutable. This would reduce the model to a partial equilibrium model in which there would be no feedback from the distribution of workers to the distribution of wages. Autor, Katz, and Kearney (2008) present a simple model in which college workers and high school workers work in separate sectors producing intermediate goods via linear production technologies. These intermediate goods are combined into a final good through use of a CES production technology. They estimate the elasticity of substitution between the two types of intermediate inputs to be 1.57. In an attempt to obtain some comparability between my work and other work using the Lucas-Island model set-up I choose to present results for $\chi = 2.25$ which is about halfway between the lower bound of Shimer and Alvarez's elasticities and that of Autor, Katz and Kearney.¹¹

For simplicity, the shape of the function, $g(\delta)$, is chosen to be uni-modal, symmetric around zero and differentiable almost everywhere. I employ the beta distribution, stretched into the interval $[-\pi, \pi]$, which typically is governed by two parameters. However, for the beta density to be symmetric the two parameters must be equal which reduces the number of parameters to be calibrated.¹² Denoting the shape parameter for the $g(\delta)$ function to be ξ , I choose ξ to be equal to 2.65. This gives the $g(\cdot)$ function a time-invariant bell-shape. The distribution for idiosyncratic fixed costs of retraining is chosen to be gamma with mean and variance denoted by μ_z and σ_z^2 and the distribution from which period shocks to θ are drawn is chosen to be a symmetric beta distribution with parameter α_ϵ .

In choosing the values of the parameters, one target that I attempt to approximately match is an average unemployment rate of 5.7% which is the sample average of the U.S. unemployment rate between the first quarter of 1948 and the first quarter of 2010. Given the parameterizations used in the simulations to follow, the ensemble average of mean unemployment rates across

¹¹I have examined models calibrated for χ between 1.75 and 10 and the qualitative results stand although the general equilibrium effects are more prominent for parameterizations near $\chi = 2$.

¹²In using the beta density to determine the shape of the function $g(\cdot)$ it is noted that $g(\delta)$ is not differentiable at exactly π and $-\pi$. However, by normalizing the value of $g(\pi) = g(-\pi) \cong 0$ the level of output per worker is so low in these occupations that nobody will be employed as the wage will be below b . Furthermore, in the numerical solution, I do not make use of derivatives in the value function iteration procedure (see Appendix).

sample runs is 5.8%. In simulations of 50000 periods, the average unemployment rate is also roughly 5.8%.

In recent work, Kambourov and Manovskii (2009a) find that between the early 1970s and the mid-1990s the average annual rate of occupational switching is 18%.¹³ They show that in the early 1970s the occupational switching rate was about 16% and rose to approximately 20% by the mid-1990s. In a similar effort, using Current Population Survey data from 1979-2006, Moscarini and Thomsson (2007) argue that at a monthly frequency, the occupational switching rate across three-digit occupational classification in the US is approximately 3.5%.¹⁴ With this number, if individuals face an independent switching probability from month-to-month, then annual mobility would be 34%. While Moscarini and Thomsson find results consistent with the upsurge in occupational mobility between 1979-1995 found by Kambourov and Manovskii, it is shown that there was a large decline in occupational mobility between 1996-2004 only to see a cyclical rebound after 2004.

I am interested in choosing parameters to fit median distances between occupation switching as well as frequency of occupation switches and so I prefer to use targets on distance between occupations and rates of occupation switching that are derived from the same data set. Recent work by Poletaev and Robinson (2008) uses information from the Dictionary of Occupational Titles (DOT) to construct distances between occupations. The Dictionary of Occupational Titles contains information that allows them to distinguish between occupations that are similar and those that are different in the underlying tasks they perform. Poletaev and Robinson (2008) use this information on job characteristics and construct low-dimensional skill vectors for each occupation in the Dictionary of Occupational Titles. They then produce measures of distance on these skill vectors to interpret distance between any two occupations. Robinson (2010) combines these measures of distance with survey data recorded in the March Current Population Surveys in order to provide a description of the patterns of occupational mobility and distance between occupations involved in switches.

Importantly, Poletaev and Robinson (2008) and Robinson (2010) distinguish between occupation switches which are determined by changes in the occupational digits as classified in the DOT and occupation switches which involve a change in the main tasks (or significant changes

¹³Kambourov and Manovskii (2009a) use 3-digit occupation classification codes employed in the U.S. Census and the PSID dataset to measure occupational mobility. A worker is defined as having switched occupations if he reports that he currently works in a different occupation than his last reported occupation in the previous year's survey. The sample is restricted to male heads of households, aged 23-61, who are not self- or dual-employed and who are not working for the government.

¹⁴Moscarini and Thomsson (2007) use the monthly CPS data from 1979-2006 and calculate occupational mobility rates for all male individuals between the ages 16-64 who were employed in consecutive months. Due to changes in the classification system of three-digit occupation codes in 1983 and due to changes in the CPS interviewing methodology in 1994, their sample is broken down into three subsamples (1979-1982, 1983-1993, and 1994-2006) within which comparisons are consistent but across which there are issues of consistency of measurement.

in the vector of skills). As they note, some occupation switches as determined by switches in the 3-digit occupational code involve negligible change in the skills involved in the job. In our model, workers who switch occupations are incurring increasing costs in the distance between their current occupation and their new occupation, which is to characterize the costs of learning new tasks inherent in occupation switching. Thus I exploit some of the empirical findings from Robinson (2010) in the baseline parameterization which document the rate of occupational switching when accounting for non-trivial changes in the skill vector associated with two different occupations.

In Robinson (2010) it is found that using 3-digit occupation code changes as the definition of occupation switches, from 1982 to 2001, the rate of occupational mobility was 9.08% in a pooled sample with a minimum annual rate of 7.06% in 2001 and a maximum annual rate of 10.85 in 1988. It is noted that while the level of measured mobility is lower than those found by Kambourov and Manovskii (2009a) and Moscarini and Thomsson (2007) the pattern of occupational mobility is similar in that there is an upward trend in occupational mobility from the early 1980s to the early 1990s and then a downward trend from the early 1990s until the early 2000s. When occupation switches are classified as requiring a change in the main skill in the skill vector with the highest score then the pooled rate of occupational mobility between 1982-2001 drops to 5.71%. In further refinements of these skill-portfolio measures of occupational mobility that use information on the level and distance of changes in the skill portfolio to classify switchers, the rate of occupational mobility falls as low as 3.58%. I use these numbers to provide a rough target of about 5.7% for annual occupational switching rates in the simulations in order to help pin down an interesting region of the parameter space.

In order to construct the occupation switching rate over a year, I count the fraction of all those employed at the beginning of the fourth quarter of a year who switched occupations at least once between the first and fourth quarters of the respective year. I include fourth quarter switchers due to time aggregation issues. In repeated simulations of 50000 quarters I find that occupation mobility rates calculated this way vary between 5.9% and 6.1%.¹⁵

I use a quadratic function to represent retraining costs to the worker $\varphi(x) = \frac{\eta}{2}x^2$. The parameters are chosen to loosely fit the median distance between occupational switches as reported by Robinson (2010). Recent work by Poletaev and Robinson (2008) uses information from the Dictionary of Occupational Titles to construct distances between occupations. Using these distance metrics, Robinson (2010) argues that, using the 3-digit occupation classifications in the

¹⁵If instead I count the occupation switching rate as the fraction of all workers who switched occupations at least once over the last year (including the unemployed) then I obtain occupations switching rates between 5.5% to 5.7% across simulations of 50000 periods. Finally, if I count the occupation switching rate as those across the population who switched at some point across the year but were employed at the time of switching (but may be unemployed in the last quarter of the year) then I again obtain switching rates between 5.4% and 5.7%. All the different methods of calculating annual switching rates hover around the target of 5.7%.

Table 2: Quarterly U.S. Data (1948Q1 - 2010Q1) vs Simulated Data

| Output per Worker | U.S. Data | Model |
|---------------------------|-----------|---------------------------|
| Standard Deviation | 0.0301 | 0.0320 (0.0253,0.0431) |
| Quarterly Autocorrelation | 0.9688 | 0.9621 (0.9453,0.9828) |

Notes: The logarithm of the data are taken and detrended using the band pass filter isolating frequencies between 6 and 200 quarters. In constructing the simulated data, the model is simulated 1000 times. The logarithm of the simulated data are taken and detrended using the band pass filter isolating frequencies between 6 and 200 quarters. The mean of the relevant statistics are presented in the table with their 95% simulation bands in parenthesis.

Dictionary of Occupational Titles, the median distance in changes for occupational switchers is approximately 0.907 with a maximum distance between occupations (for a 4-dimensional skill vector measure of distance) of 8.072. Relative to our model, where the maximum distance between two occupations is π , I target a median distance normalized by π of approximately 0.11. In our simulations I calculate the median distance in occupational switches for all movers from the first quarter of a year to the first quarter of the next year. Doing so with the parameter values displayed in Table 1 yields an average normalized distance across occupational switchers over 50000 periods of 0.095.

In order to help give some context to the level of consumption enjoyed during unemployment, $b(\psi)$ is 33% of the average wage. This puts unemployment benefits paid in the model towards the lower end of replacement ratios offered in the U.S. Finally, I choose the parameters so as to approximately fit the volatility and persistence of average labour productivity as measured by output per worker in the U.S. data. Here persistence is measured as the correlation between average labour productivity and its first lag.

Table 2 provides statistics on average labour productivity for the U.S. economy after filtering out high frequency movements as well as similarly treated simulation data from the model. The parameterization chosen for the model yields average labour productivity dynamics that resemble those of the U.S. economy over the sample period. The volatility of average labour productivity of the U.S. economy falls in the 95% simulation intervals as does the observed first order autocorrelation. On average, this parameterization yields slightly less persistence in labour productivity than observed in the data. It is also instructive to note that the 95% band on the volatility of average labour productivity from the model is quite wide indicating that there is excessive volatility in the model. This will work against the model in terms of obtaining simulated time series within which there are episodes characterized by uninterrupted periods

of either positive or negative growth in specific variables. Note that outside of targeting some statistics that may be thought of as “level” statistics, these are the only statistics involving volatility and persistence that are directly targeted by the parameterization.

3.2 The Dynamics Within Episodes

In this subsection, I first provide an example of a simulation in which the dynamics of the model economy qualitatively resembles that of the U.S. economy. Then the model’s dynamics are examined to highlight conditions on the distribution of workers across occupations and on the sequence of reallocation shocks which are conducive to generating the types of dynamics in occupation switching rates and wage inequality as observed in the U.S. data. Such conditions are constructed and fed back into the model to see how frequently the model qualitatively replicates observed joint dynamics in the data when provided with favourable conditions.

3.2.1 An Example

The example presented here is an attempt to highlight mechanisms at play in the model. The model generates episodes that are qualitatively similar to those observed in the U.S. economy between the mid-1970s and the early-2000s. Specifically, the example consists of an episode in which wage inequality rises in the top and bottom halves of the distribution along with the occupation switching rate and this is followed by an episode in which there is polarization in the wage distribution along with a falling occupation switching rate. For the rest of the paper, I will refer to the first type of event as an “Episode 1” event while the second type of event will be referred to as an “Episode 2” event. Finally, any periods which are not of Episode 1 or Episode 2 types I will refer to as belonging to “the Rest”.

Figure 4 shows the annualized occupation switching rates and measures of wage inequality from a subsample of a sample run with 249 periods, which is the same sample size as used in constructing quarterly statistics from the U.S. economy. In this example, I drew a sequence of 249 shocks and then replaced a window of 80 periods worth of shocks with those illustrated in the upper-left panel of Figure 4. I then filtered all the simulated data to isolate medium run frequencies. I display only the results from the 80 periods with the shock sequence which I purposefully constructed to highlight the main mechanism. The bottom-left panel shows that the trend in the annual rate of occupation switching rises between first nine years and then falls over the rest of the sample period.¹⁶ The bottom-right panel shows that the trends in the log 50/10 and log 90/50 ratios of the annual wages constructed by averaging quarterly wages. Note that both ratios rise continuously for the first seven years of the sample while their trends jointly

¹⁶The ragged lines in Figure 4 are the simulated data and the smooth lines are the trends from the filtered data.

rise for the first nine years. Wage polarization occurs between years eleven and seventeen.

The upper-right panel illustrates the average annual wages of the 90th, 50th and 10th percentile wage earners. It can be seen that in levels, the wage at the bottom of the distribution falls faster than the median wage during the first eight years. Furthermore, the wage of the 90th percentile earner rises over this period and, on average, the median wage exhibits downward movement so we observe rising wage inequality at both ends of the wage distribution. Subsequently, between years eleven and seventeen, there is a rise in the wage of the 10th percentile wage earner relative to the median wage while the wage of the 90th percentile wage earner rises relative to the wage of the median wage earner resulting in polarization in the wage distribution.

In Figure 5 the Episode 1 and 2 events are examined in more detail. The figure presents the wage distribution and the distribution of individuals across occupations for the quarterly simulated data during the episodes with the panels in the top row illustrating characteristics of the Episode 1 event and the bottom panels illustrating the characteristics of the Episode 2 event. The left panels plots the distribution of workers across occupations (relative to the frontier) in the first and last quarter of the episode.¹⁷ In the middle panels, a bird's eye view of the distribution of workers across occupations is presented for each period of the episode. The top line with the square markers plots the occupation of the 90th percentile individual when occupations are ordered on the $[-\pi, \pi]$ interval. The bottom line with the square markers plots the occupation of the 10th percentile individual when ordered on the $[-\pi, \pi]$ interval. By similar construction the top line with the circle markers represents the 75th percentile individual, the solid line shows the location of the 50th percentile individual and the lower line with circle markers shows the occupation of the 25th percentile worker. The right panel shows the wage of the 90th, 75th, 50th, 25th and 10th percentile wage earner, respectively, as time progresses.

As illustrated by the solid line in the left panel of the top row, at the beginning of the episode, workers are concentrated in occupations near the productivity frontier and the mode of the distribution of workers across occupations is near $\delta = 0.75$. From the middle panel of the top row, it can be seen that over the episode, the occupation of the median worker drifts from about $\delta = 0.3$ to approximately $\delta = 1$. Given that, on average the annual occupation switching rate is around 5.7%, this implies that the cause of this drift is a sequence of negative shocks as shown in the upper-left panel of Figure 4. As the occupation of the 50th percentile individual (as ordered on the $[-\pi, \pi]$ interval) moves towards π it means 50% of individuals are being housed by a fewer measure of occupations. These occupations are relatively less productive and the increase in supply of individuals in those occupations depresses the wage resulting in a downward trend in the wage of the 25th and 10th percentile wage earner. On the other hand, as the occupation of the 50th percentile individual moves towards π , this also means that the

¹⁷I discuss the distribution of workers across occupations (by occupation name) later in the paper.

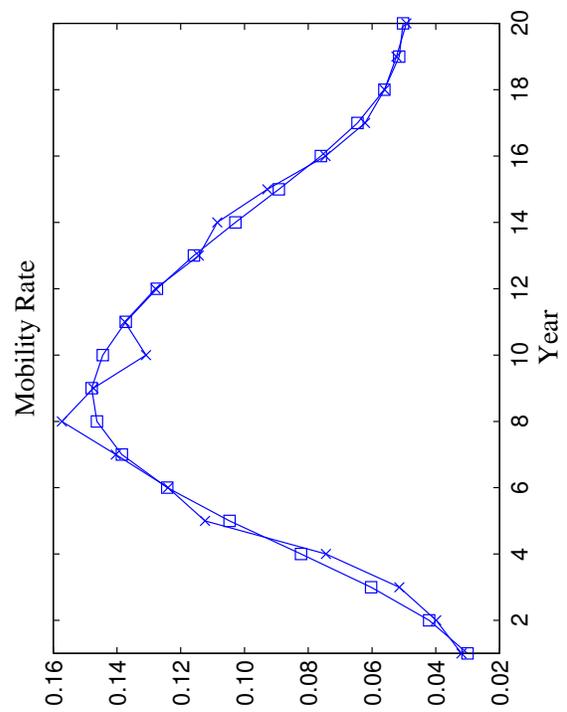
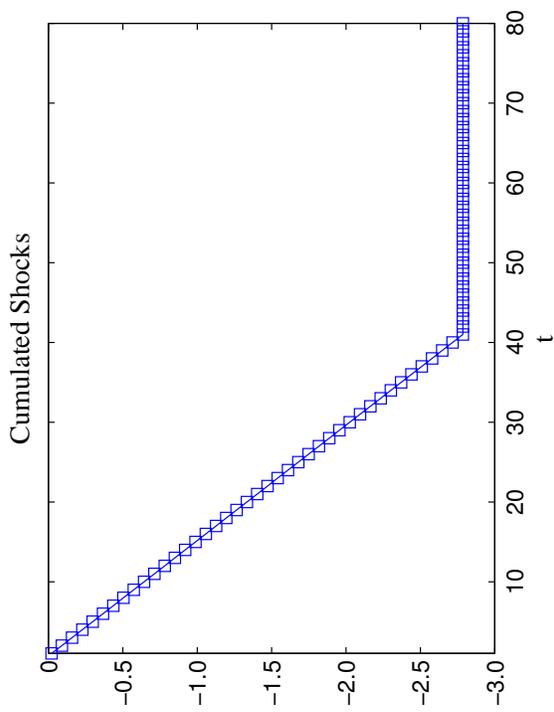
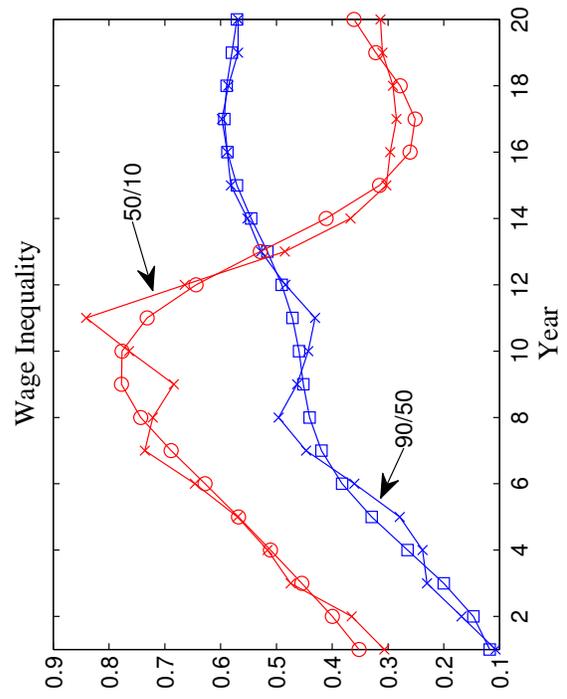
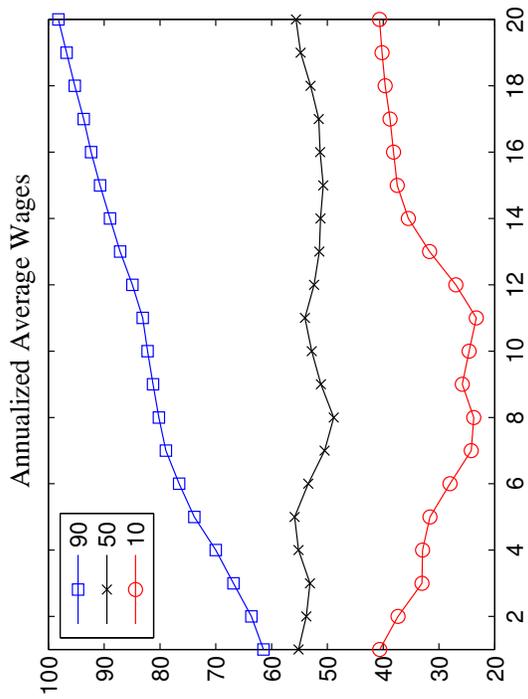


Figure 4: Simulated Sample Run

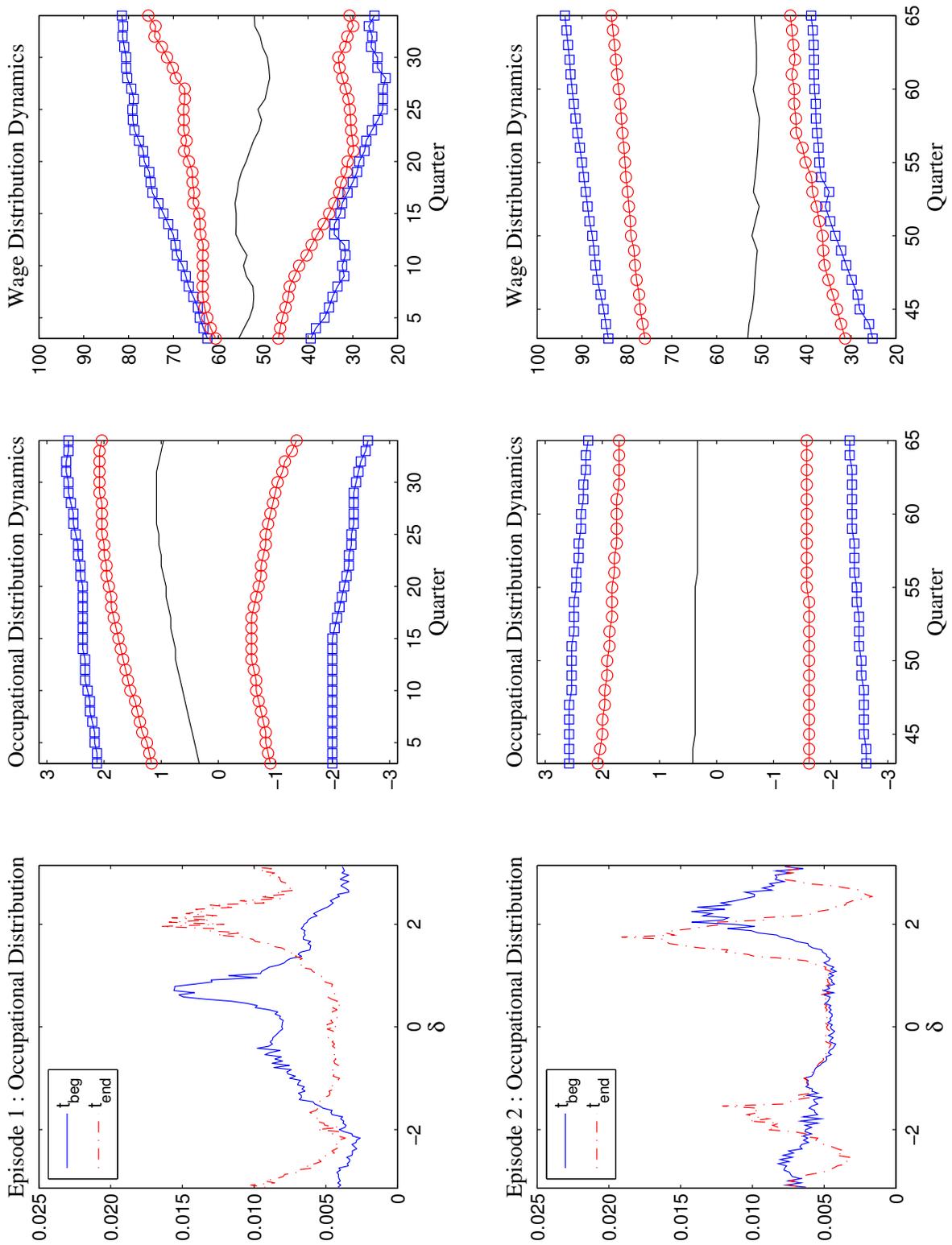


Figure 5: Occupational Distribution and Wage Distribution Dynamics During Episodes

remaining 50% of individuals are now spread out over a larger set of occupations. This puts upward pressure on their wages. Also, as a large fraction of workers experience a reduction in productivity (and more workers become unemployed) aggregate output falls putting downward pressure on the wages of all workers. By the end of the episode, a large mass of workers is housed in a small set of unproductive occupations, namely those a distance between $[2, \pi]$ away from the most productive occupation and those near $-\pi$ away from the frontier. This is shown by the broken line in the upper-left panel.

The upper-right panel reveals that as the productivity frontier drifts, the wages of the 90th percentile rises consistently and, while the median wage fluctuates, it does not exhibit any notable trend. Note that the difference between the median earner's wage and those of the 90th, 75th, 25th and 10th percentile earners is increasing over this episode which gives rise to the increase in wage inequality both in the top and bottom halves of the wage distribution. It is important to note that many workers are pushed into very unproductive occupations and choose to become rest unemployed. This results in an increase in the unemployment rate during this episode and the measure of employed workers from which the wage distribution is constructed is decreasing.¹⁸ Again, note that as the measure of workers in the most productive occupations decrease, the wage paid to these workers increases relative to the situation when there is a large measure of workers working in the most productive occupation. During this episode, this reduction in supply of workers in the most productive occupations puts more upward pressure on their wages than does the general equilibrium effect that a reduction in aggregate output puts on their wages.

Simultaneous to the rise in wage inequality there is an increase in the occupation switching rate. This is because workers who experience a decline in the wage choose to switch occupations in an effort to re-obtain higher wages. As these workers chase higher wages, the occupation switching rate rises but wage inequality rises because some workers do not draw low enough fixed costs of switching and remain stuck in the lower wage occupations.

In contrast to the start of the Episode 1 event when workers are bunched near the most productive occupation, at the start of the Episode 2 event workers are more spread out across occupations with more workers attached to relatively unproductive occupations. This is shown by the solid line in the lower-left panel. By construction the shocks during this episode are forced to be zero. A striking feature of this episode is that in contrast to that of Episode 1 there is relatively little drift in the occupation of the median worker between the first and last periods. As seen in the middle panel of the bottom row, the reallocation of workers across

¹⁸Given that unemployment benefits are a fraction of the average wage, as the average wage falls, unemployment consumption also falls tempering the fluctuation in the unemployment rate. Simulations not displayed in the paper show that with fixed unemployment benefits, the model exhibits much more volatility although the main mechanisms at play do not change.

occupations during the episode is mainly due to workers out in the tails of the distribution choosing to switch to more productive occupations. This is clear as the occupations of the 90th, 75th, 25th and 10th percentile individuals all move towards zero over the episode. As more individuals move into their preferred occupations over the episode, in the absence of long sequence of shocks in any direction, fewer workers face the incentive to switch occupations. This results in the fall in the occupation switching rate over the episode.

The property that workers near the most productive occupations choose not to switch is due to the assumption on the shape of the function governing productivity across occupations and the convex costs of switching occupations. Once a worker is housed in an occupation near the productivity frontier, the marginal gain in wages due to increased productivity becomes very low (the function $g(\delta)$ is quite flat near $g(0)$). Coupled with the convexity of retraining costs the results is that, even in the absence of fixed costs, the marginal cost of retraining exceeds the marginal benefits of switching at occupations near the productivity frontier. This is why workers tend to settle into mid-productivity occupations during the episode and is why there is a bi-modal distribution of workers across occupations at the end of Episode 2.¹⁹

Focusing on the wage dynamics, it is useful to recall that the wage of the i^{th} percentile wage earner (i) rises as the level of productivity in the occupation housing the i^{th} percentile wage earner increases, (ii) falls as the measure of workers attached to the occupation of the i^{th} percentile wage earner increases, and (iii) rises as aggregate output increases. While workers at the bottom end of the wage distribution are moving to higher productivity occupations and earning higher wages, they are also bunching up in these higher paying wages which reduces the wage inequality in the bottom half of the wage distribution. At the bottom end of the distribution, the rise in wages associated with increased productivity outstrips the downward pressure of having more workers attached to the occupations in the lower end of the wage distribution. Over the episode, the median wage does not change much so that the increase in productivity attached to the occupation of the median wage earner is mostly offset by the increase in the supply of earners attached to the median wage occupation. Note that as workers get pulled out of unemployment, the identity of the median wage earner changes as can the occupation attached to this median earner. Finally, there is a notable rise in the wages paid at the top end of the wage distribution. As there is little change in the distribution of workers in occupations near the productivity frontier during this episode, most of the gains in wages at the top arise because aggregate output is increasing during this period.

¹⁹Simulations using the calibrated model shows that if the economy is hit with zero shocks for hundreds of periods then the distribution of workers across occupations becomes bi-modal with large mass at mid-productivity occupations. This is also observed in the partial equilibrium model, discussed below, when intermediate goods are perfect substitutes in the production of the final consumption good.

3.2.2 Characteristics of the Episodes

In order to further refine the understanding of the different types of episodes, properties of the sequence of shocks as well as the distribution of workers across occupations that are typical of each type of episode are extracted from an ensemble of simulations. I simulate the model 1000 times with each sample run consisting of 249 quarters (after dropping 500 burn-in periods). The data from each sample run is then filtered to leave behind medium-run trends. Finally, the filtered data from each sample run is scanned for each type of episode. Once all the sample runs are scanned, and all instances of each episode is saved, each type of episode is examined in an attempt to characterize distinguishing features of the distribution of workers across occupations during the episode as well as the sequence of shocks hitting the economy during the episode.

Table 3 displays some properties of Episode 1 and 2 events and also displays statistics for periods in which wage inequality properties of the Episodes are met but occupational mobility rate properties are not. The title “Only Rise” refers to episodes in which the log 90/50 and the log 50/10 ratios simultaneously trend upwards but occupation switching rates do not trend up for the required duration. In similar spirit, the title “Only Polarization” refers to periods during which the log 90/50 ratio trends upwards while the log 50/10 ratio trends downwards but the occupational switching rate does not trend downwards for a sufficiently long duration. In constructing this table, a variable trends in a particular direction if the annual growth rate of the variable’s trend is strictly positive for at least 6 consecutive years or strictly negative for at least 6 consecutive years. I discuss the implications of varying this cut-off later in the paper. Note that this table is just describing properties within episodes but does not assume that Episode 1 and 2 events occur consecutively. In constructing the statistics, workers are ordered by their occupations on the $[-\pi, \pi]$ interval. The distance of the i^{th} percentile worker as ordered on $[-\pi, \pi]$ is denoted by δ_i . All lengths are normalized so that any differences in the table can be interpreted as a percentage of the interval $[-\pi, \pi]$.

The first two rows of Table 3 give some information about the size of shocks during particular types of episodes. We can see that the average size of shocks are not statistically different from those that occur during non-episodic periods. The second row shows the average of the maximum cumulated sum of shocks during a particular episode. This is meant to capture the feature that a sequence of strictly positive shocks or a sequence of strictly negative shocks is one of the most important ingredients of episodes featuring swings in occupation switching rates. We can see that periods exhibiting properties of Episode 1 tend to exhibit sequences of shocks that cause the mode of the labour productivity function to move in one direction much farther than those periods that are characterized by the properties of Episode 2. In order for the occupation switching rate to exhibit an upward trend for a prolonged length of time, a

Table 3: Examining the Episodes : The Shocks and The Distribution

| The Shocks | Only Rise | Episode 1 | Only Polarization | Episode 2 | Rest |
|---|---------------------|--------------------|--------------------|--------------------|--------------------|
| $\mu(\epsilon_t)$ | 0.0005 (0.0120) | 0.0002 (0.0124) | 0.0005 (0.0063) | 0.0004 (0.0054) | 0.0000 (0.0022) |
| $\max(\sum_{t=1}^i \epsilon_t), 1 \leq i \leq T$ | 0.3163 (0.1118) | 0.3256 (0.1103) | 0.1950 (0.0880) | 0.1734 (0.0764) | - - |
| The Distribution (Whole Episode) | | | | | |
| $\mu(\delta_{60} - \delta_{40})$ | 0.1511 (0.0185) | 0.1509 (0.0169) | 0.1823 (0.0369) | 0.1994 (0.0316) | 0.1686 (0.0112) |
| $\mu(\delta_{75} - \delta_{25})$ | 0.3904 (0.0273) | 0.3904 (0.0250) | 0.4390 (0.0575) | 0.4523 (0.0634) | 0.4128 (0.0151) |
| $\mu(\delta_{90} - \delta_{10})$ | 0.6884 (0.0313) | 0.6912 (0.0296) | 0.7282 (0.0629) | 0.7247 (0.0710) | 0.6995 (0.0123) |
| The Beginning and the End | | | | | |
| $\mu_1(\delta_{60} - \delta_{40})$ | 0.1427 (0.0235) | 0.1422 (0.0228) | 0.1799 (0.0396) | 0.1997 (0.0324) | - - |
| $\mu_T(\delta_{60} - \delta_{40})$ | 0.1706 (0.0422) | 0.1732 (0.0421) | 0.1882 (0.0428) | 0.1974 (0.0384) | - - |
| $\mu_1(\delta_{75} - \delta_{25})$ | 0.3655 (0.0235) | 0.3635 (0.0214) | 0.4361 (0.0641) | 0.4614 (0.0679) | - - |
| $\mu_T(\delta_{75} - \delta_{25})$ | 0.4445 (0.07539) | 0.4481 (0.0733) | 0.4391 (0.0596) | 0.4385 (0.0609) | - - |
| $\mu_1(\delta_{90} - \delta_{10})$ | 0.6536 (0.0267) | 0.6528 (0.0270) | 0.7367 (0.0705) | 0.7453 (0.0759) | - - |
| $\mu_T(\delta_{90} - \delta_{10})$ | 0.7608 (0.0633) | 0.7677 (0.0577) | 0.7101 (0.0636) | 0.6989 (0.0660) | - - |
| $\mu_1(\mathcal{M}(\delta) - \mathbb{M}(\delta))$ | 0.1126 (0.0668) | 0.1096 (0.0656) | 0.2097 (0.1112) | 0.2497 (0.1010) | - - |
| $\mu_T(\mathcal{M}(\delta) - \mathbb{M}(\delta))$ | 0.1999 (0.1022) | 0.2085 (0.0947) | 0.1901 (0.0697) | 0.1978 (0.0703) | - - |

Notes: To construct the data the model is simulated 1000 times. For each episode that occurs we record the statistics above. Then the mean and the standard deviation of the statistic across all simulations in the ensemble is calculated. $\mu(\delta_x - \delta_y)$ indicates the mean of the $\delta_x - \delta_y$ interval across an episode. $\mu_1(\delta_x - \delta_y)$ indicates the mean of the $\delta_x - \delta_y$ interval for the first four quarters of a particular episode and $\mu_T(\delta_x - \delta_y)$ indicates the mean of the $\delta_x - \delta_y$ interval for the last four quarters of an episode. Denote the mean of the distribution of workers across occupations by $\mathcal{M}(\delta)$ and the mode by $\mathbb{M}(\delta)$. Then $|\mathcal{M}(\delta) - \mathbb{M}(\delta)|$ is the absolute value between the mean of the distribution of occupations of workers and the mode of the distribution of workers across occupations. Denote by $\mu_1(|\mathcal{M}(\delta) - \mathbb{M}(\delta)|)$ the ensemble average of this statistic when $|\mathcal{M}(\delta) - \mathbb{M}(\delta)|$ is annualized for the first year of each episode in the ensemble. The statistic $\mu_T(|\mathcal{M}(\delta) - \mathbb{M}(\delta)|)$ is similarly constructed but for the last year of each episode in the ensemble. The mean and standard deviation across all episodes of a particular type is reported. All level statistics are expressed as percentages of the interval $[-\pi, \pi]$.

sequence of potentially large shocks in the same direction tends to push a large mass of workers away from their optimal occupation providing the incentives for a large measure of workers to retrain. Note that while the average maximum of cumulated shocks for a Polarization episode is smaller than that of an Episode 1 run, the Episode 2 value is even smaller than that of a Polarization episode indicating that indeed the property of small sequences of cumulated shocks in one direction comes from the need to have a falling occupation switching rate.

The third through fifth rows display average properties of the distribution of individuals across the entirety of each type of episode. While the average distribution for episodes in which only wage inequality is rising or when the occupation switching rate trends upwards along with wage inequality does not differ much from normal times, the distribution of individuals during times characteristic of Episode 2 features much more dispersion. It is important to note that during any interval in time, workers may be constantly retraining so the distribution of individuals may change drastically within an episode. In order to understand whether these distributions vary within episodes, the last six rows of Table 3 presents the mean of various statistics as measured in the first and last years of each episode across the ensemble.

The first statistic to note is that the ensemble mean of the $\delta_{75} - \delta_{25}$ interval is significantly smaller in the first period of Episode 1 events relative to non-episodic periods (0.3635 on average relative to 0.4128). This means that the set of occupations held by those between the 25th and 75th percentile workers ordered on the $[-\pi, \pi]$ interval is small at the beginning of Episode 1 events so that there is much equality at the beginning of these episodes. If the economy is hit with a sequence of shocks mostly in the same direction, it is likely that a large measure of workers will become disadvantaged in the sense of a fall in their labour productivity. As time progresses some of these workers will retrain to higher paying occupations which leads to an increase in wage inequality at the top half. As we see in the row for $\mu_T(\delta_{75} - \delta_{25})$, in the last period of Episode 1 events the distance covered by the $\delta_{75} - \delta_{25}$ interval tends to be significantly larger than in the first period (0.4481 relative to 0.3635). This statistic backs the story that given the average sequence of shocks hitting the economy during Episode 1 events, workers in these occupations tend to switch occupations spreading out the measure of occupations covered by the middle 50 percent of workers as ordered by occupational distance from the most productive occupation. Similar characteristics are featured by the $\delta_{60} - \delta_{40}$ interval.

Looking at the interval covered by the $\delta_{90} - \delta_{10}$ difference, it can be seen that, on average, this interval is larger at the end of Episode 1 events than during the average non-episodic periods (0.7677 compared to 0.6995). This indicates that at the end of Episode 1 events many workers are clustered in the occupations with low levels of labour productivity. Low labour productivity coupled with a large supply of workers in those occupations result in low wages and, likely, high levels of unemployment.

Summarizing, for episodes in which wage inequality rises both at the top and bottom along with rising occupation switching rates, it is found that in the early periods of these episodes, there is much clustering of workers in occupations near the most productive occupation. Then a sequence of large technological innovations occurs which push the most productive occupation away from the mass of workers. This results in incentives for this mass of workers to retrain. As it takes time for workers to switch to their preferred occupations due to fixed and convex costs of occupation switching, there is a rise in wage inequality and the occupation switching rate as workers slowly reallocate over occupations.

On the flip-side, the average of the first period $\delta_{75} - \delta_{25}$ spread is larger in Episode 2 events (0.4614) as compared to Episode 1 events (0.3635). Also, we see that while the $\delta_{90} - \delta_{10}$ spread covers a larger area in the first period than the average non-episodic period, by the end of the episode this spread is not very different from the average non-episodic period. The intuition here is that with much dispersion of individuals across occupations and few shocks, individuals will tend to switch occupations early in the episode and by the end of the episode fewer workers are interested in incurring further costs to switch occupations. This gives rise to the fall in the occupation switching rate over the episode.

Lastly, the statistic $\mu(|\mathcal{M}(\delta) - \mathbb{M}(\delta)|)$ measures the absolute value of the distance between the mean of the distribution of workers across occupations and the mode of this distribution. This statistic is meant to be informative about the location of the mode at the beginning and end of episodic events. Typically the mean of the distribution tends to oscillate around zero. Note from Table 3, that at the beginning of episodes in which wage inequality rises both at the top and bottom of the distribution, the mode tends to be relatively close to the mean. In contrast at the beginning of episodes in which wage polarization occurs, the mode of the distribution tends to be significantly further away from the mean. It is also informative to note that the absolute value of the distance between the mode and the mean tends to be insignificantly different between the end of an Episode 1 event, and the beginning of an Episode 2 event.

Summarizing the characteristics of Episode 2 events, there is typically a large dispersion of workers around the most productive occupation in the early going and, relative to Episode 1 events, the drift in the most productive occupation is less pronounced during the episodes than their Episode 1 counterparts. This lack of large drift gives rise to a decrease in the occupation switching rate within the episode. Furthermore, at the beginning of Episode 2 events, the mode of the distribution of workers tends to be significantly further away from the production frontier than at the beginning of Episode 1 events. This is reflected in larger measures of wage inequality at the beginning of Episode 2 events. As more workers are clustered in less productive occupations, this gives rise to individual incentives to reallocate towards more productive, higher

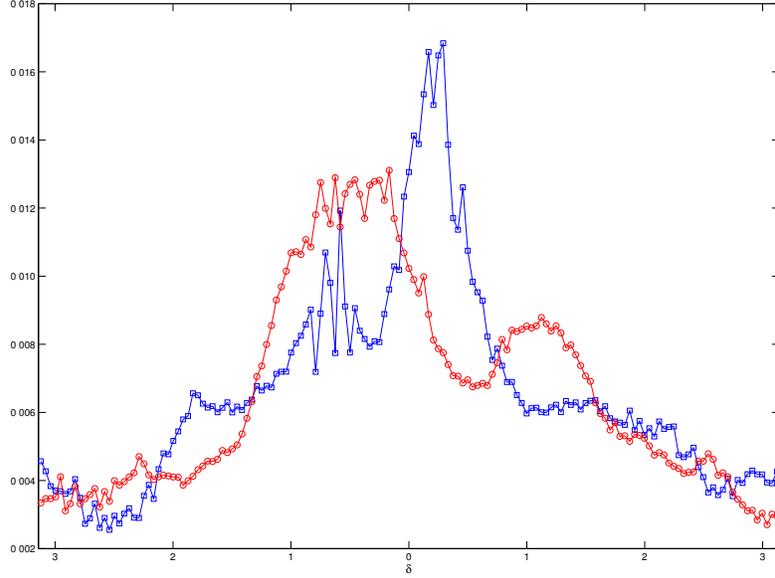


Figure 6: Examples of Initial Distributions

paying occupations in the absence of further reallocation shocks. I now use these observations in order to construct distributions and shock sequences which will give rise to episodes that, qualitatively, are similar to those of the U.S. experience.

3.3 Constructing Episodic Events

In order to test the conditions required for the model to qualitatively produce Episode 1 and 2 events in succession, I construct an ensemble of distributions that possess particular characteristics. I focus on initial distributions such that the following characteristics are satisfied : (i) $\frac{\delta_{60}-\delta_{40}}{2\pi} \in (0.10, 0.12)$, (ii) the normalized interval $\frac{\delta_{75}-\delta_{25}}{2\pi} \in (0.33, 0.34)$, (iii) the normalized interval $\frac{\delta_{90}-\delta_{10}}{2\pi} \in (0.65, 0.66)$, (iv) the absolute deviation of the mode of the distribution from the mean, $|\mathcal{M} - \mathbb{M}(\delta)| < 0.05$, and (v) the normalized absolute value of the mean, $|\mathcal{M}(\delta)| < 0.01$. With the exception of the interval $\frac{\delta_{90}-\delta_{10}}{2\pi}$ this puts all other measures below their mean by one standard deviation at the beginning of Episode 1 events as in Table 3. The requirement that the absolute value of the mean of the distribution be less than 0.01 come from the statistic $\mu_1(|\mathcal{M}(\delta)|)$ constructed from the ensemble of distributions as in the other statistics from Table 3. Figure 6 provides examples of distributions which satisfy these requirements.

In order to construct such distributions, I simulated the model for 500 burn-in periods, then continued the run for an additional 249 periods looking for distributions satisfying the

aforementioned criterion. I saved every simulation run which carried a period featuring a desired distribution until I had 100 successful simulation runs.

Next, in order to construct episodic events, for each of these 100 stored runs I then ran a simulation using the same sequence of shocks as the original run up until the period in which the distribution of workers satisfied the above criterion. At this point I hit the economy with sequence of identical shocks lasting 40 periods such that the absolute value of the maximum distance moved by the most productive occupation was either $\frac{2\pi}{3}$ or $0.44 \cdot 2\pi$. Furthermore, during these 40 periods, the most productive occupation was required to move away from the original location of the mode of the distribution of workers. This sequence of shocks was an attempt to generate an Episode 1 event. For the next 40 periods, I shut down the shocks in order to generate an Episode 2 event. After these 80 periods, I hit the economy with the rest of the sequence of stored shocks. In other words, for each of the 100 stored runs, I ran a simulation in which I hit the economy with a sequence of stored shocks but where 80 periods of the original sequence of shocks were replaced with 40 consecutive periods of “extreme” shocks and then 40 periods of zero shocks. For each of the 100 sample runs, the simulated data was then filtered. I then checked to see if this window of 80 periods was able to generate Episode 1 and 2 events. The example in subsection 3.2.1 was constructed in the manner just described.

Table 4 displays the outcome of these attempts to construct qualitative episodes as observed in the U.S. data. The table shows the average number of episodic events across 100 simulations. In the top panel of Table 4, the most productive occupation rotated constantly two-thirds the way around the occupation ring over 40 quarters and then did not rotate at all for the next 40 quarters. This is approximately the mean shift of the identity of the most productive occupation across Episode 1 events as shown in Table 3. In the middle and bottom panels, the corresponding shift is 44% of the way around the occupation ring which is approximately one standard deviation away from the mean shift across typical Episode 1 events as shown in Table 3. For this table, in defining an event as having occurred, I varied the number of consecutive years in which the requirements for an episode needed to be met. Whereas in Table 3, an episode was registered when the conditions for the event were met for six consecutive periods, here the number of years is varied between four and seven. There is no rule for determining the length of the stylized episodes and empirical work has uncovered one of each type of episodes given available data. This table serves to highlight the speed at which workers reallocate over occupations in the model, specifically when no shocks are hitting the economy.

The top panel filters the simulated data using the band pass filter to isolate medium-run frequencies. It is clear that as the number of years required to register an event increases, the ability of the model to generate episodes is reduced. In the table, the row “Only Rise” displays the frequency with which the requirement of Episode 1 events that the trend in occupational

Table 4: Fraction of Simulations Featuring Stylized Facts of Fixed Duration

| | | | | |
|-------------------|-------|-------|-------|-------|
| Exercise #1 (BP) | 4 Yrs | 5 Yrs | 6 Yrs | 7 Yrs |
| Only Rise | 77% | 61% | 43% | 23% |
| Episode 1 | 61% | 45% | 24% | 17% |
| Only Polarization | 78% | 42% | 17% | 6% |
| Episode 2 | 34% | 15% | 5% | 4% |
| Both | 24% | 8% | 3% | 2% |
| Exercise #2 (BP) | 4 Yrs | 5 Yrs | 6 Yrs | 7 Yrs |
| Only Rise | 99% | 90% | 78% | 57% |
| Episode 1 | 99% | 90% | 74% | 48% |
| Only Polarization | 100% | 99% | 82% | 51% |
| Episode 2 | 96% | 95% | 69% | 38% |
| Both | 95% | 86% | 55% | 22% |
| Exercise #3 (BP) | 4 Yrs | 5 Yrs | 6 Yrs | 7 Yrs |
| Only Rise | 75% | 61% | 43% | 24% |
| Episode 1 | 74% | 58% | 42% | 21% |
| Only Polarization | 100% | 98% | 83% | 65% |
| Episode 2 | 95% | 82% | 74% | 57% |
| Both | 69% | 50% | 35% | 12% |
| Exercise #2 (HP) | 4 Yrs | 5 Yrs | 6 Yrs | 7 Yrs |
| Only Rise | 100% | 98% | 92% | 79% |
| Episode 1 | 99% | 98% | 92% | 77% |
| Only Polarization | 100% | 100% | 95% | 84% |
| Episode 2 | 100% | 99% | 87% | 63% |
| Both | 99% | 97% | 81% | 55% |

Note : The table presents simulated annual statistics on the trends of occupational mobility rates and wage inequality measures. In order to construct these statistics, the annual averages of their quarterly counterparts are constructed and filtered leaving behind only frequencies longer than 8 years; this is to capture medium- to long-run behaviour.

mobility rates continually rises fails, but the requirement that wage inequality continually rises is met. Contrasting this row with the row displaying the frequency of Episode 1 events shows that when the economy is buffered by shocks which continually shift the productivity frontier, as the duration required to register an episode is lengthened, the less likely it appears that increased wage inequality is accompanied by persistent increases in the occupational mobility rate.

It is also clear from the top panel that as the duration required to register an episode increases, the likelihood of wage polarization drops of rapidly, by a factor of 13 between four and seven years. Also, by contrasting the third and fourth rows of the top panel, it is clear that when the productivity frontier shifts constantly and cumulatively by $\frac{2\pi}{3}$ over ten years, wage

polarization is only accompanied by persistently declining occupational mobility rates about a third to a half of the time.

In the second panel, the exercise is repeated but with a cumulative shift in the productivity frontier of 44% of the distance around the occupation ring. This panel highlights the effects of the “speed of retraining”. Now, the model does not seem to have much difficulty generating episodes of either type. Notably, as the duration required in order to register an event rises from four to seven years, the frequency of episodes drops by a factor of about two or four. Thus it appears that in order to generate the kind of dynamics observed in the U.S. economy, the model would suggest that the period between the late-1970s and late-1980s was a particularly turbulent time in terms of the drastic change in demand for occupational specific tasks. On the other hand, the model also suggests that the period during which wage polarization and the drop in occupational mobility rates occurred may have simply been a period during which there was not much change in occupational task demands and the results was just a matter of workers reallocating themselves optimally across occupations.

The third panel repeats the exercise keeping the shocks the same as in the second panel but allowing for the initial distribution of workers across occupations to be more dispersed than in the top two panels of the table. Specifically, the initial distribution of workers are constructed such that the statistics characterizing the shape of the distribution of workers across occupations are between the mean value of the statistics (for the beginning of the period from Table 3) and the mean of the statistic less one standard deviation. By contrasting the second and third panels, we can see that there is not much loss in the ability of the model to generate Episode 2 events but the model has more problems generating Episode 1 events. In other words, given a persistent sequence of technological innovations which favour a set of occupations that are dissimilar to the initially productive occupations, compressing the initial distribution of workers around the most productive occupation increases the likelihood of observing episodes in which wage inequality rises both at the top and bottom ends of the wage distribution and occupational mobility rises for a reasonably long period of time.

Recalling from Figure 2 that use of the HP-filter with the standard weight in handling annual data, produces much smoother trends in wage inequality, Exercise #2 of Table 4 uses the HP-filter on the same simulations as in the second panel. Using the HP-filter produces episodes with increased frequency relative to use of the band pass filter. In this exercise it can be seen that using the HP-filter increases the likelihood of successfully registering longer duration events which means that much of the movements feature a much slower moving component than is left behind by the medium-run bandpass filter specification.

Summarizing, the model is able to produce episodes that qualitatively line up with the U.S. experience, when the initial distribution of workers across occupations is compressed around the

Table 5: Fraction of Simulations Featuring Stylized Facts

| | Only Rise | Episode 1 | Only Polarization | Episode 2 | Both |
|---------------------|-----------|-----------|-------------------|-----------|-------|
| 4 Years | 89.3% | 86.7% | 90.5% | 63.1% | 58.3% |
| 5 Years | 77.6% | 74.2% | 71.9% | 40.4% | 35.7% |
| 6 Years | 59.7% | 53.7% | 43.5% | 20.4% | 15.3% |
| 7 Years | 34.9% | 28.1% | 19.3% | 8.6% | 3.6% |
| Partial Equilibrium | | | | | |
| 4 Years | 94.9% | 93.0% | 78.6% | 37.7% | 34.1% |
| 5 Years | 88.2% | 85.0% | 51.2% | 16.3% | 13.0% |
| 6 Years | 65.3% | 57.7% | 21.9% | 5.5% | 3.5% |
| 7 Years | 34.3% | 27.7% | 9.1% | 0.9% | 0.2% |

Note : In order to construct the statistics above, the relevant annual growth rates of the trends are calculated. “Only Rise” is the case in which at least one n -year period features consecutive non-negative growth in both the log 50/10 ratio, and the log 90/10 ratios. “Only Rise & O.M.” is the case of “Only Rise” in addition to the requirement that the occupational switching rate grows in each of the n -years%. “Only Polarization” is the case in which the log 90/50 ratio exhibits positive growth, and the log 50/10 ratio exhibits negative growth in each year of an n -year period. “Polarization & O.M.” is the case, that in addition to the conditions of “Only Polarization”, occupational mobility rates decrease for each year in the n -year period. Finally, “All” is the case in which there is at least one n -year episode in the sample of 249 quarters in which “Rise & O.M.” occurs and another in which “Polarization & O.M.” occur. Note that the table presents simulated annual statistics on the trends of occupational mobility rates and wage inequality measures. In order to construct these statistics, the annual averages of their quarterly counterparts are constructed and filtered leaving behind only frequencies longer than 8 years; this is to capture medium- to long-run behaviour.

most productive occupation and the economy is subjected to a persistent sequence of innovations which favour a set of occupations that use tasks which are quite dissimilar to the set of tasks initially favoured by technologies.

3.3.1 Frequency of Events

Instead of creating conditions under which episodes are likely to arise and examining the model’s behaviour, Table 5 presents the frequency of episodes in standard simulations. As in Table 4, the features of any particular type of episode must run uninterrupted on annual simulated data for at least n years in order to register an episode where n is varied from four to seven. In constructing the table, the model is simulated for 1000 runs, each of 249 quarters. Annual data is constructed by taking quarterly averages, the data is filtered and then each simulation is scanned for episodic events. The total number of episodes of each type is counted across the 1000 simulation runs. The table displays the fraction of simulations containing at least one episode of the specified type. The column “Both” refers to the fraction of simulations containing at least one Episode 1 event and at least one Episode 2 event.

Given the particular calibration, it is difficult to obtain both Episode 1 and Episode 2 events in a single simulation when an event requires episodic characteristics to occur uninterrupted for at least six years. This is due to the excessive volatility in the model. Given that parameterization of the model requires workers to move relatively frequently and, on average, to move to occupations using dissimilar tasks, when workers move they tend to move a fair distance around the occupational ring. Without sufficiently long periods of time in which the economy is hit with shocks that continually favour a particular set of occupations, it is unlikely to observe both Episode 1 and Episode 2 events in a single run.

In order to understand the role that the distribution of workers plays on the frequency of episodes, Table 5 also presents the frequency of episodes for an ensemble produced by a partial equilibrium version of the model. The partial equilibrium case is such that there is perfect substitutability between intermediate inputs in the production of the final good. In this case, the wage paid by any occupation is equal to a worker's labour productivity. In the partial equilibrium case, parameters are recalibrated to match the targeted statistics. Furthermore, unemployment benefits are fixed (not a fraction of the average wage) but calibrated such that on average, across a run of 10000 periods, unemployment benefits equal about 33% of the average wage. Preventing the distribution of workers across occupations from affecting the level of unemployment benefits keeps the exercise as close as possible to the notion of partial equilibrium.

As can be seen, the partial equilibrium model does not produce sample runs that contain at least one of each Episode 1 and 2 type events more frequently than the baseline calibrated model. However, the reason for this is that the partial equilibrium case generates Episode 1 events as frequently as the baseline while generating Episode 2 events much less frequently. This implies that the general equilibrium forces in the model are important for the possibility of polarization in the wage distribution. As the distribution of workers across occupations does not affect wages, it is difficult to get workers to spread themselves out far enough across occupations and even more difficult to get the period of reallocation to be drawn out over a long period of time; workers chase the high productivity occupations and converge on them quickly.

The difficulty of the model in generating episodes of longer length is not surprising as there is no saving technology for individuals to insure themselves against occupational risk.²⁰ Furthermore, the presence of fixed costs of retraining and the fact that workers cannot work in the period in which they switch occupations gives workers an incentive to move far around the

²⁰Hawkins and del Rio (2012) examine the effects of allowing individuals who face borrowing constraints to save in order to insure themselves against occupational risk in a Lucas-Prescott island model. Their paper examines the steady state effects of occupational risk on aggregate output and savings while also detailing the impacts on individual mobility decisions.

occupational ring each time an occupational switch occurs.

4 Discussion

The driving force behind occupational switching in this paper has been shocks to the relative labour productivity across occupations over time. There is nothing preventing the driving force to be shocks to relative demand for the output produced by occupations. For example, if the CES aggregator used by the final goods sector featured weights on the intermediate inputs produced across occupations, relative shocks to these weights would produce shifts in the wages paid by occupations holding constant the distribution of workers across occupations. This would give individual workers incentives to switch occupations resulting in behaviour similar to the model above. This paper uses shocks to relative labour productivity as a driving force because a large body of research examining the rising wage inequality in the U.S. economy during the 1980s and 1990s has focused on innovations to computer technologies. More recent discussions about wage inequality have mentioned that offshoring of tasks performed by particular sets of occupations may be driving down the demand for workers attached to these occupations. The mechanics of this model suggest that the results would be similar to having innovations in relative demand across occupational output.

A recent body of work in line with Autor, Levy, and Murnane (2003) and Acemoglu and Autor (2010) has investigated the hollowing out of the task distribution. These papers have noted that between the 1970s and the mid- to late-2000s there was a decrease in the measure of workers employed in occupations that call on routine manual tasks. Autor, Levy, and Murnane (2003) argued that the rising accessibility of computer inputs reduced the demand for workers to perform routine manual tasks because computers could be used as substitutes. On the other hand, computers appear to be complimentary inputs for workers performing non-routine cognitive tasks. Other papers such as Acemoglu and Autor (2010) and Cortes (2011) have extended this notion of non-routine tasks to include low-paying non-routine manual tasks. This has led to the finding that there was a hollowing out of employment in occupations that are highly associated with routine manual tasks with a corresponding increase in employment across occupations requiring non-routine cognitive and non-routine manual tasks.

Figure 7 returns to the example simulation from Section 3.2.1. The figures in the top row plot the measure of workers across occupations but the horizontal axis now plots the identity of the occupation (not the distance from the productivity frontier). The plots track the measure of workers attached to particular occupations as time passes. The solid line depicts the distribution of workers at the beginning of the Episode 1 event. One interpretation is that there is a large measure of workers attached to occupations calling for routine-manual tasks; occupations $i \in [5, 2\pi]$ and $i \in [0, 1]$. Over time, computer innovations make non-routine tasks

(those require by occupations $i \in [1, 5]$) in greater demand. By the end of the 20 year period, the distribution of workers across occupations is depicted by the dashed line in the upper-right figure. Notice that workers have mostly shifted away from those occupations that rely on manual tasks. Some of these workers have shifted towards more productive occupations which can be thought of as requiring non-routine cognitive tasks while a larger measure are now employed in lower-productivity occupations, say those occupations $i \in [1, 2]$. These can be thought of as non-routine manual tasks. Given this interpretation, the model can also provide a dynamic framework to consider the joint dynamics of task-demand, occupation mobility rates and measures of wage inequality.

An outstanding issue with the choice to model occupations as being located around an occupation ring is that the model does not tend to have occupations which are systematically higher paying occupations; ex ante all occupations are just as likely to be the most productive occupation. This ignores issues such as high-paying, white-collar occupations versus low-paying blue-collar occupations. Pushing this argument further, the model does not currently allow for the possibility that particular characteristics are attached to particular occupations. This means that workers who are endowed with particular abilities are not restricted to be attached to a particular subset of jobs. Thus while it is possible to make some occupations more likely to be the most productive over time, the current version of the model does not cement individuals to certain types of occupations. It would be useful for future extensions of the model to allow occupations as having some fixed characteristics as well as restricting individuals to certain subsets of occupations, however this model was constructed to permit a first pass examination of macroeconomic issues and therefore left some microeconomic details aside.

Jaimovich and Siu (2012) recently examined the task polarization phenomenon and note that most of the polarization of tasks between 1990 and 2010 has occurred during small windows surrounding the three recessions of the period. A difference between the data examined in this paper is that this paper is focused on the joint dynamics of wage inequality and occupational mobility rates. In line with recent work on task mobility, Jaimovich and Siu (2012) partition occupations into the same three categorizations as Acemoglu and Autor (2010). Thus, while not directly comparable, the spirit of this paper and that of Jaimovich and Siu (2012) is certainly similar.

Lastly, I note that the model in this paper is closely related to the “canonical” model (as dubbed by Acemoglu and Autor (2010)) used in much work studying wage inequality but also shares some properties of the “Ricardian” models featured in Costinot and Vogel (2010) and Acemoglu and Autor (2010) which is an extension of the canonical model. In the canonical model, the supply side of production mirrors that of the model in this paper. The difference is that the canonical model is typically a static model upon which comparative statics are

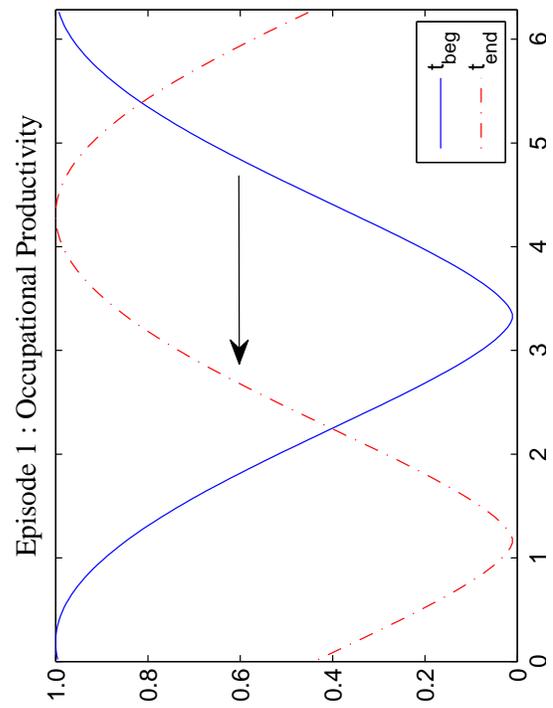
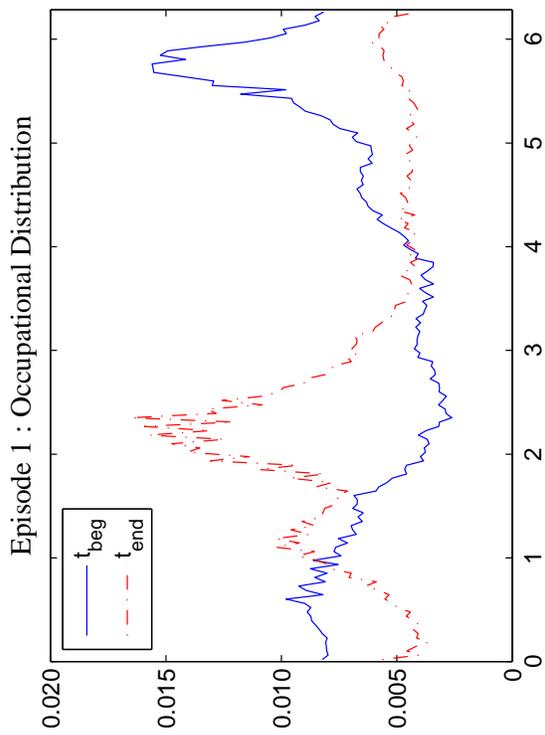
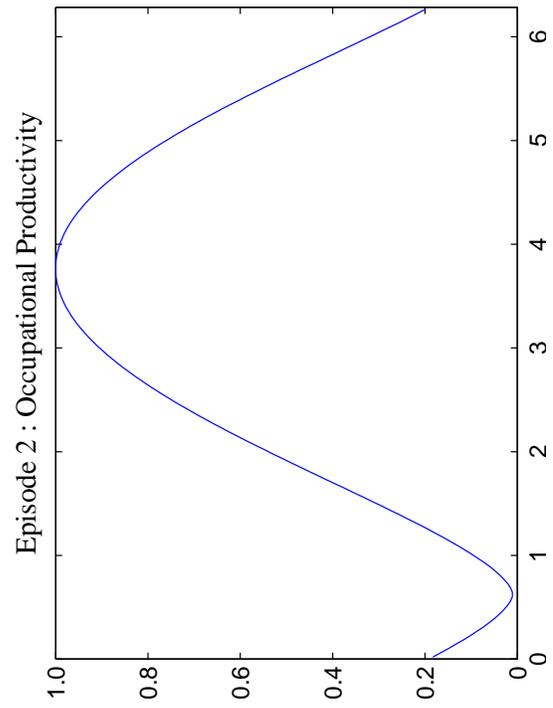
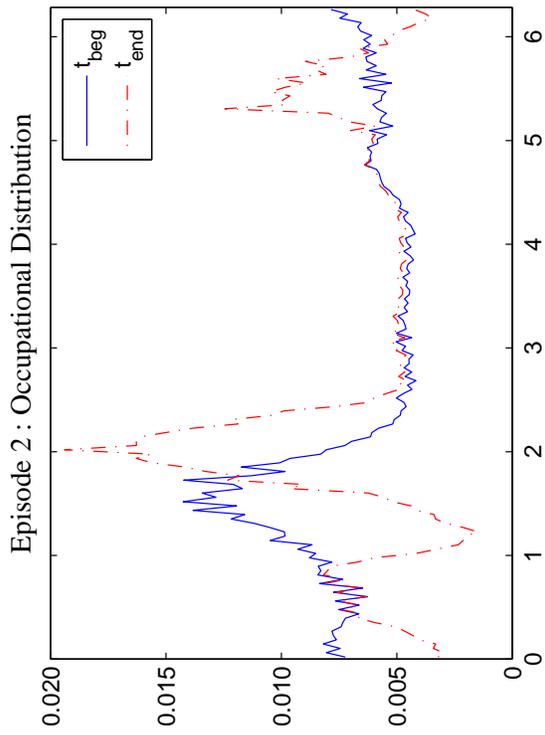


Figure 7: The Reallocation of Tasks

performed. In order to close the model workers are assigned to particular occupations (or education class, sector, etc.). In the Ricardian model, workers are endowed with skills of different types (as in a standard Roy model) and these skills are transferable across tasks. Output is generated by occupations which require particular tasks. The appealing feature of the Ricardian model is that it allows workers to choose how much of their skill endowments to allocate across different tasks used in production. As individuals differ in their vector of skill endowments, individuals possess comparative advantages between tasks and as productivity or demand for different tasks change over time, the wage distribution along with the distribution of workers' labour effort across tasks change.

In my model, while workers only possess a single skill at any given point in time, they are able to endogenously vary this skill over time which gives rise to a feature similar to that of the Ricardian model. In particular, individuals are malleable across occupations or tasks over time. However, the Ricardian model endows individuals with a particular set of skills which is time invariant so some individuals will never perform certain tasks as well as others. This is clearly a plausible assumption that I sidestep for tractability with the benefit of being able to simulate the dynamics over time.

5 Conclusion

This paper develops a dynamic model of equilibrium occupational choice in which an economy is subjected to aggregate reallocation shocks and workers may choose to incur costs to retrain in order to move to higher wage occupations. The model provides a framework to understand the joint patterns in the occupational mobility rate and wage inequality as observed in the U.S. economy. Specifically, the model produces periods in which either (i) occupational mobility rises along with increasing wage inequality in both the top and bottom halves of the wage distribution or (ii) occupational mobility falls, while a rise in inequality in the top half of the wage distribution is accompanied by a fall in inequality in the bottom half. Within this model economy, the changing pattern in observed wage inequality and occupational mobility can be thought about as arising simply from the distribution of workers across occupations and the set of occupations favoured by innovations to demand for occupational output or by innovations to labour productivity which favour some occupations over others.

One objective in developing this model is to provide some insight into the design of social insurance packages providing consumption insurance to the unemployed as well as retraining subsidies to workers who possess skills for occupations whose output experiences a reduction in demand. As an equilibrium problem this is still beyond the reach of this model for tractability reasons and is left for future work.

Appendix

A Solution Method

The solution method used in this paper is similar to the methods proposed in Reiter (2002) and Reiter (2010).

A.1 Outline of the Algorithm

In order to approximate the equilibrium distribution of workers across occupations, I use a grid of moments \mathcal{M} which is formed as a Cartesian product from the set of one-dimensional grids for a set of m moments, \mathcal{M}_j with $j \in \{1, \dots, m\}$. In this paper I used two moments : the mean and the variance of the location of workers across a set of bins in the interval $[-\pi, \pi]$. Let \hat{m} denote an element of \mathcal{M} . I partitioned the interval $[-\pi, \pi]$ into n_δ bins so that the density of workers can be approximated by a histogram following the procedure of Young (2010). Denote the set of mid-points of the bins by $\Delta = \{\delta_1, \dots, \delta_{n_\delta}\}$.

1. For each grid point in $\hat{m} \in \mathcal{M}$ guess at the histogram $\psi(\delta, \hat{m})$, for all $\delta \in \Delta$.
 - (a) For each grid point $\hat{m} \in \mathcal{M}$ guess at the transition function $\hat{\Gamma} : [-\pi, \pi] \times \mathbb{R}_+ \rightarrow [-\pi, \pi] \times \mathbb{R}_+$ where, as I use the mean and variance, I have $\hat{m} \in [-\pi, \pi] \times \mathbb{R}_+$. This provides us with a guess $\hat{m}' = \Gamma(\hat{m})$ for each $\hat{m} \in \mathcal{M}$.
 - i. Taking the functions $\psi(\delta, \hat{m})$ and $\hat{m}' = \Gamma(\hat{m})$ as given solve for the individual decision rules using the value function iteration procedure discussed below. Denote the optimal decision rule for choosing the employment state by $\iota(\delta, \hat{m})$ and the decision rule for retraining by $x(\delta, \hat{m})$.
 - (b) For each grid point $\hat{m} \in \mathcal{M}$, using the decision rules $\iota(\delta, \hat{m})$ and $x(\delta, \hat{m})$ along with the histogram $\psi(\delta, \hat{m})$, calculate the \tilde{m}' consistent with individual optimization and the conjectured equilibrium distribution. Check to see if $\tilde{m}' = \hat{m}'$. If so, I have found the individual rules and the equilibrium transition function that are consistent with each other taking the equilibrium distribution as given.
2. Taking the transition function and decision rules from above, simulate the aggregate economy for T periods saving the distribution of workers for each $t = 1, \dots, T$. Now using the procedure below, approximate the equilibrium density function $\tilde{\psi}(\delta, \hat{m})$. For each $\hat{m} \in \mathcal{M}$ check to see if the distance between $\psi(\delta, \hat{m})$ and $\tilde{\psi}(\delta, \hat{m})$ is sufficiently small. If so, then stop. Otherwise, repeat until convergence is achieved.

A.2 Value Function Iteration

In order to solve the individual worker’s dynamic programming problem, I take as given the transition function for the moments as well as the distribution of workers for each $\hat{m} \in \mathcal{M}$. The value functions for the workers have three arguments, δ , μ and σ^2 . I solve the dynamic programming problem with a slight alteration of the method proposed by Barillas and Fernández-Villaverde (2007). I do not use information on the derivative of the value functions as is standard in the endogenous grid point method because of problems that potential kinks in the continuation values of employment, unemployment or retraining pose along with the discrete choice nature of the worker’s problem. Particularly, by using derivatives of the continuation value, due to kinks, it is possible that two occupations, say $\hat{\delta}' = \delta + \hat{x}$ and $\tilde{\delta}' = \delta + \tilde{x}$ are optimal continuation occupations from the same current occupation, δ even though, in equilibrium, one of the occupations may not be an optimal occupation to move to starting from δ . The assumption, by using the derivative of the continuation occupation, in solving backwards as in the endogenous grid point method (EGM) of Carroll (2006) and Barillas and Fernández-Villaverde (2007) is that every occupation in the continuation grid of occupations can be reached in equilibrium from some occupation in the grid of current occupations; an assumption that is not necessarily true in equilibrium.

Thus, instead of exploiting derivatives, I use the following procedure which combines some ideas of traditional value function iteration procedures and some ideas from the EGM methods. Particularly, I solve the dynamic programming problem forwards (in contrast to backwards as in the EGM) which avoids the problem detailed in the previous paragraph. I follow the EGM method in that I reduce the number of times that I solve for the expected continuation values per iteration of the value functions which reduces computational time immensely.²¹

I now describe the procedure. Define a grid on the continuation occupation $G' \equiv \{\delta'_1, \dots, \delta'_{N'}\}$ and a grid for current occupations $G \equiv \{\delta_1, \dots, \delta_N\}$. Construct a large (symmetric) grid for values of x , $X \equiv \{x_1, \dots, x_{N_x}\}$ which contains values between $-\pi$ and π . The algorithm to iterate on the value functions is

1. Start with $k = 0$ and guess values for $V^k(\delta, \mu, \sigma^2)$, $U^k(\delta, \mu, \sigma^2)$, and $T^k(\delta, \mu, \sigma^2)$.
2. Using Gaussian quadratures, for each $(\hat{\delta}, \mu, \sigma^2)$, where $\hat{\delta}$ denotes the end of period occupation for a worker, construct the expected continuation values

²¹As a check of this procedure, I solved Model 1 of Barillas and Fernández-Villaverde (2007) and found that the solutions were virtually identical but my procedures was slightly slower. However, my procedure was able to solve certain cases in my model which ran into difficulties when using derivatives of the continuation value functions.

$$W^k(\hat{\delta}, \mu, \sigma^2) = \int \int \max \left\{ T^k(\delta', \mu', \sigma^{2'}), U^k(\delta', \mu', \sigma^{2'}), V^k(\delta', \mu', \sigma^{2'}) \right\} dF(\delta'|\hat{\delta})dH(z')$$

3. Construct a grid $G(\delta) = \{\delta + x_1, \dots, \delta + x_{N_x}\}$ for each δ and then interpolate the vector $W(\hat{\delta}, \mu, \sigma^2)$ across this grid. This yields the expected continuation value for each end of the period occupation in the grid $G(\delta)$. For each $\delta \in G$ construct the functions

$$\begin{aligned} V^{k+1}(\delta, \mu, \sigma^2) &= u(w(\delta, \mu, \sigma^2)) + \beta W^k(\delta, \mu, \sigma^2) \\ U^{k+1}(\delta, \mu, \sigma^2) &= u(b) + \beta W^k(\delta, \mu, \sigma^2) \\ T^{k+1}(\delta, \mu, \sigma^2) &= \max_{x \in X} \left\{ u(b) - \varphi(x) + \beta W^k(\delta + x, \mu, \sigma^2) \right\}. \end{aligned}$$

4. If

$$\sup_{i,j,N} \left\{ \frac{\left\| \begin{bmatrix} V^{k+1}(\delta_N, \mu_i, \sigma_j^2) - V^k(\delta_N, \mu_i, \sigma_j^2) \\ U^{k+1}(\delta_N, \mu_i, \sigma_j^2) - U^k(\delta_N, \mu_i, \sigma_j^2) \\ T^{k+1}(\delta_N, \mu_i, \sigma_j^2) - T^k(\delta_N, \mu_i, \sigma_j^2) \end{bmatrix} \right\|}{1 + \sup_{i,j,N} \left\| \begin{bmatrix} V^k(\delta_N, \mu_i, \sigma_j^2) \\ U^k(\delta_N, \mu_i, \sigma_j^2) \\ T^k(\delta_N, \mu_i, \sigma_j^2) \end{bmatrix} \right\|} \right\} \geq 1.0e^{-6}$$

then $k \rightsquigarrow k + 1$ and go to 2. Otherwise, stop and store the value functions, $V(\delta, \mu, \sigma^2)$, $U(\delta, \mu, \sigma^2)$, $T(\delta, \mu, \sigma^2)$ and the decision rule $x(\delta, \mu, \sigma^2)$.

For the results shown in the paper, I used a linear grid for X with 501 points and for G and G' I used a linear grid with 151 points. In the partial equilibrium model, increasing these grid sizes did not yield significantly different results for the simulation output.

A.3 Simulating the Economy

In order to simulate the economy I use a method similar to that proposed by Young (2010). Specifically, I use a histogram over a fixed grid of the occupations to construct moments for the next period. Suppose I break the interval $[-\pi, \pi]$ into N_B equal length bins whose midpoints serve as the “location” of the bin. Once I know the measure of workers in each bin, I can easily calculate the moments desired. In simulating the model for T periods I then use the following procedure :

1. Take the histogram of workers at the beginning of the period t as given.
2. Given this histogram, construct the mean and variance of the distribution. Denote the period t mean by μ_t and the variance by σ_t^2 .

3. Use the value functions as obtained from the value function iteration and interpolate across the moments in \mathcal{M} to obtain $V(\delta, \mu_t, \sigma_t^2)$, $U(\delta, \mu_t, \sigma_t^2)$, $T(\delta, \mu_t, \sigma_t^2)$ and $x(\delta, \mu_t, \sigma_t^2)$ for all $\delta \in G$.
4. Using these value functions, for each bin, determine the cut-off cost (from the distribution of idiosyncratic fixed cost of retraining) for which workers would prefer to retrain rather than be employed or unemployed. Also determine, for each bin, if workers are not to retrain whether they would prefer to be employed or unemployed.
5. For those who elect to retrain, determine the bin in which their new occupation falls. Assign all workers who retrain to their new bins.
6. Draw an aggregate shock for period $t + 1$. Adjust the bins for all workers given the end of period t histogram and the period $t + 1$ reallocation shock.
7. For each bin, let a fraction $1 - \beta$ of the workers withdraw from the labour force. In each bin add a measure $\frac{1-\beta}{N_B}$ of new workers.
8. If $t = T$ then stop. Otherwise, $t \rightsquigarrow t + 1$ and go to 1.

A.4 The Transition Function for the Moments

In order to determine the transition function for the moments I follow the steps below. Suppose N_ϵ nodes are used to evaluate the Gaussian quadrature in the distribution for the reallocation shocks, that there are N_μ elements in the grid for the mean and that there are N_{σ^2} elements in the grid for the variance.

1. For possible triple $(\mu_i, \sigma_j^2, \epsilon_k)$, conjecture the continuation moment pair $\{\mu'_{ijk}, \sigma'^2_{ijk}\}$.
2. Using these continuation moments, for each $\{\mu, \sigma^2\} \in \mathcal{M}$, construct the continuation values

$$W(\delta, \mu, \sigma^2) = \int \int \max \left\{ T(\delta', \mu', \sigma'^2), U(\delta', \mu', \sigma'^2), V(\delta', \mu', \sigma'^2) \right\} dF(\delta'|\delta) dH(z')$$

and then construct the functions

$$\begin{aligned} \hat{V}(\delta, \mu, \sigma^2) &= u(w(\delta, \mu, \sigma^2)) + \beta W(\delta, \mu, \sigma^2) \\ \hat{U}(\delta, \mu, \sigma^2) &= u(b) + \beta W(\delta, \mu, \sigma^2) \\ \hat{T}(\delta, \mu, \sigma^2) &= \max_{x \in X} \left\{ u(b) - \varphi(x) + \beta W(\delta + x, \mu, \sigma^2) \right\}. \end{aligned}$$

3. For each $\{\mu, \sigma^2\} \in \mathcal{M}$, simulate the economy for one period and find the realized moment pairs $\{\tilde{\mu}'_{ijk}, \tilde{\sigma}'_{ijk}\}$. Note that this is a *realized* pair meaning one pair for each possible shock $\epsilon' \in \{\epsilon_1, \dots, \epsilon_{N_\epsilon}\}$.
4. For some set of weights $\{\omega_{i,j,k}\}$, if

$$\sup \left\{ \frac{\left| \omega_{i,j,k} \begin{bmatrix} \tilde{\mu}'_{ijk} - \mu'_{ijk} \\ \tilde{\sigma}'_{ijk} - \sigma'_{ijk} \end{bmatrix} \right|}{1 + \sup \left| \omega_{i,j,k} \begin{bmatrix} \tilde{\mu}'_{ijk} - \mu'_{ijk} \\ \tilde{\sigma}'_{ijk} - \sigma'_{ijk} \end{bmatrix} \right|} \right\} \geq 1.0e^{-4}$$

then set $[\mu'_{ijk}, \sigma'_{ijk}] = \lambda[\mu'_{ijk}, \sigma'_{ijk}] + (1 - \lambda)[\tilde{\mu}'_{ijk}, \tilde{\sigma}'_{ijk}]$, where $\lambda \in [0, 1)$ and go to 2. Otherwise, stop.

In this process I introduce a set of weights $\{\omega_{i,j,k}\}$ in order to reduce the computation time. Effectively, the weights may be chosen to down-weight the importance of nodes that are unlikely to be reached in equilibrium, for example, for mean and variances that are at the extremes of their respective grids or for reallocation shocks that are of very low probability of being drawn.

A.5 The Equilibrium Distribution

In order to characterize the equilibrium distribution necessary to determine the equilibrium wages at each triplet (δ, μ, σ^2) for each $\delta \in G$, $\mu \in \mathcal{M}_1$ and $\sigma^2 \in \mathcal{M}_2$, I follow a procedure set forth by Reiter (2002) and Reiter (2010). I repeat his procedure here. First I construct what is referred to as a “reference distribution” for each $\{\mu, \sigma^2\} \in \mathcal{M}_1 \times \mathcal{M}_2$.

1. Initialize the reference distribution $\psi^R(\delta, \mu, \sigma^2)$ either by a uniform distribution or by some other means.²²
2. Starting with this reference distribution, simulate the economy for $T_{Burn} + T$ periods and then discard the statistics from the first T_{Burn} “burn-in” periods. In the results shown, I used $T_{Burn} = 2000$ and $T = 25000$. Store the mean, variance and histogram from each of these T periods. Let $\psi_{sim,t}$ denote the simulated histogram from period $t = 1, \dots, T$ (it is a vector). Let $\mu_{sim,t}$ and $\sigma_{sim,t}^2$ be the corresponding mean and variance from period t of the stored simulation data.

²²In the computations, I simulate the partial equilibrium model in order to obtain some realizations for equilibrium distribution given a set of parameters. Then I solved our general equilibrium model many times while reducing the elasticity parameter χ each time. This was the most time consuming process in calibrating the model.

3. Update the $\psi^R(\delta, \mu, \sigma^2)$ by the weighted sum

$$\begin{aligned}\psi^R(\delta, \mu, \sigma^2) &\equiv \sum_{t=1}^T \gamma_t \psi_{sim,t} \\ \gamma_t &\equiv \frac{d([\mu, \sigma^2], [\mu_{sim,t}, \sigma_{sim,t}^2])^\nu}{\sum_{t=1}^T d([\mu, \sigma^2], [\mu_{sim,t}, \sigma_{sim,t}^2])^\nu}\end{aligned}$$

where, in our application, $d(\cdot)$ was the Euclidean metric and the weight ν was chosen to be -4 .

Given the reference distributions, $\psi^R(\delta, \mu, \sigma^2)$ for each pair $(\mu, \sigma^2) \in \mathcal{M}_1 \times \mathcal{M}_2$, I then constructed the “proxy distribution”. Importantly, note that the mean and variance of the reference distributions are not equal to μ and σ^2 , respectively. In order to adjust each of these reference distributions I used the following steps. For each (μ_i, σ_j^2) pair, for $i = 1, \dots, N_\mu$, $j = 1, \dots, N_{\sigma^2}$, I calculated the mean and variance of the reference distribution $\psi(\delta, \mu_i, \sigma_j^2)$. Denote this reference pair as $(\mu_{R,i}, \sigma_{R,j}^2)$.

As I used a histogram to approximate the continuous distribution, I use the notation ψ_l to denote the l^{th} bin in the histogram with $l = 1, \dots, N_\delta$. Then I solved the following constrained optimization problem for each $(\mu, \sigma^2) \in \mathcal{M}$.

$$\min_{\{\psi_l\}_{l=1}^{N_\delta}} \sum_{l=1}^{N_\delta} (\psi_l - \psi_l^R)^2$$

subject to the constraints $\psi_l \geq \frac{1-\beta}{N_\delta}$ and

$$\begin{bmatrix} \delta_1 & \delta_2 & \dots & \delta_{N_\delta} \\ \delta_1^2 & \delta_2^2 & \dots & \delta_{N_\delta}^2 \\ 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{N_\delta} \end{bmatrix} = \begin{bmatrix} \mu \\ \sigma^2 + \mu^2 \\ 1 \end{bmatrix}.$$

Once this step is complete I have a set of proxy distributions $\psi(\delta, \mu, \sigma^2)$ which I use as approximations of the equilibrium distributions.

A.5.1 Updating the Proxy Distributions

In order to update the proxy distributions so that they are consistent with equilibrium behaviour, I guessed a set of proxy distributions $\psi(\delta, \mu, \sigma^2)$ for each pair $(\mu, \sigma^2) \in \mathcal{M}$ to be the equilibrium distributions. Using these guesses I solve for the individual value functions and decision rules as well as the transition function for the aggregate moments. With the individual decision rules in hand, I simulate the economy for $T_0 + T$ periods and then drop the data from the first T_0 periods. Using the histograms from the simulation I then construct updates

for the proxy distribution $\tilde{\psi}(\delta, \mu, \sigma^2)$. I then construct deciles from each of the original proxy distribution, denote these by $D(\mu, \sigma^2)$ and also from the updated proxy distributions, denoted by $\tilde{D}(\mu, \sigma^2)$. Then for some set of weights $\{\omega_{i,j}\}$, $i = 1, \dots, N_\mu$ and $j = 1, \dots, N_{\sigma^2}$, if

$$\sup \left\{ \frac{|\omega_{i,j} \otimes [D(\mu_i, \sigma_j^2) - \tilde{D}(\mu_i, \sigma_j^2)]|}{1 + \sup |\omega_{i,j} \otimes [D(\mu_i, \sigma_j^2) - \tilde{D}(\mu_i, \sigma_j^2)]|} \right\} \geq 1.0e^{-3}$$

set $\psi(\mu, \delta, \sigma^2) = \tilde{\psi}(\mu, \delta, \sigma^2)$ and repeat, otherwise stop. In the results presented I used $T_0 = 300$ and $T = 5000$.

References

- ACEMOGLU, D., AND D. AUTOR (2010): “Skills, Tasks and Technologies: Implications for Employment and Earnings,” in *Handbook of Labor Economics*, vol. 4.
- ALVAREZ, F., AND R. SHIMER (2009a): “Search and Rest Unemployment,” *Econometrica*, Forthcoming.
- (2009b): “Unemployment and Human Capital,” Mimeo, University of Chicago.
- AUTOR, D., L. KATZ, AND M. KEARNEY (2008): “Trends in U.S. Wage Inequality: Revising the Revisionists,” *Review of Economics and Statistics*, 90(2), 300–323.
- AUTOR, D., F. LEVY, AND R. MURNANE (2003): “The Skill Content of Recent Technological Change: An Empirical Exploration,” *Quarterly Journal of Economics*, 116(4), 1279–1333.
- BARILLAS, F., AND J. FERNÁNDEZ-VILLAVARDE (2007): “A generalization of the endogenous grid method,” *Journal of Economic Dynamics and Control*, 31, 2698–2712.
- BURDETT, K., S. SHI, AND R. WRIGHT (2001): “Pricing and Matching with Frictions,” *Journal of Political Economy*, 109(5), 1060–1085.
- CARRILLO-TUDELA, C., AND L. VISSCHERS (2011): “Unemployment and Endogenous Reallocation over the Business Cycle,” Mimeo.
- CARROLL, C. (2006): “The method of endogenous gridpoints for solving dynamic stochastic optimization problems,” *Economics Letters*, 91, 312–320.
- CORTES, G. M. (2011): “Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data,” Mimeo.
- COSTINOT, A., AND J. VOGEL (2010): “Matching and Inequality in the World Economy,” *Journal of Political Economy*, 118(4), 747–786.

- FIRPO, S., N. FORTIN, AND T. LEMIEUX (2010): “Occupational Tasks and Changes in the Wage Structure,” Mimeo.
- GARIN, J., M. PRIES, AND E. SIMS (2011): “Reallocation and the Changing Nature of Economic Fluctuations,” Mimeo.
- GATHMANN, C., AND U. SCHONBERG (2010): “How General is Human Capital? A Task-Based Approach,” *Journal of Labor Economics*, 28(1), 1–49.
- GAUTIER, P. A., C. N. TUELINGS, AND A. V. VUUREN (2010): “On-the-Job Search, Mismatch and Efficiency,” *Review of Economic Studies*, 77, 245–272.
- GOUGE, R., AND I. KING (1997): “A Competitive Theory of Employment Dynamics,” *Review of Economic Studies*, 64(1), 1–22.
- HAWKINS, W., AND J. M. DEL RIO (2012): “Financial Frictions and Occupational Mobility,” Mimeo.
- JAIMOVICH, N., AND H. SIU (2012): “The Trend is the Cycle: Job Polarization and Jobless Recoveries,” Mimeo.
- KAMBOUROV, G., AND I. MANOVSKII (2009a): “Occupational Mobility and Wage Inequality,” *Review of Economic Studies*, 76, 731–759.
- (2009b): “Occupational Specificity of Human Capital,” *International Economic Review*, 50(1), 63–115.
- LEMIEUX, T. (2010): “What Do We Really Know About Changes in Wage Inequality?,” in *Labor in the New Economy*, ed. by K. Abraham, J. Spletzer, and M. Harper, pp. 17–59. University of Chicago Press.
- LUCAS, R. E., AND E. C. PRESCOTT (1974): “Equilibrium Search and Unemployment,” *Journal of Economic Theory*, 7, 188–209.
- MARIMON, R., AND F. ZILIBOTTI (1999): “Unemployment vs. Mismatch of Talent: Reconsidering Unemployment Benefits,” *The Economic Journal*, 109, 266–291.
- MOEN, E. (1997): “Competitive Search Equilibrium,” *Journal of Political Economy*, 105(2), 385–411.
- MOSCARINI, G., AND K. THOMSSON (2007): “Occupational and Job Mobility in the US,” *Scandinavian Journal of Economics*, 109(4), 807–836.

- PETERS, M. (2000): “Limits of Exact Equilibria for Capacity Constrained Sellers with Costly Search,” *Journal of Economic Theory*, 95, 139–168.
- PISSARIDES, C. A. (2001): *Equilibrium Unemployment Theory*. Cambridge, MA : MIT Press, 2nd edn.
- POLETAEV, M., AND C. ROBINSON (2008): “Human Capital Specificity : Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984-2000,” *Journal of Labor Economics*, 26(3), 387–420.
- REITER, M. (2002): “Recursive computation of heterogeneous agent models,” Mimeo.
- (2010): “Solving the incomplete markets model with aggregate uncertainty by backward induction,” *Journal of Economic Dynamics and Control*, 34, 28–35.
- ROBINSON, C. (2010): “Occupational Mobility, Occupational Distance and Basic Skills : Evidence from Job-Based Skill Measures,” Mimeo.
- SHIMER, R. (2005): “The Assignment of Workers to Jobs In an Economy with Coordinantion Frictions,” *Journal of Political Economy*, 113(5), 996–1025.
- VERACIERTO, M. (2008): “On the cyclical behavior of employment, unemployment and labor force participation,” *Journal of Monetary Economics*, 55, 1143–1157.
- YOUNG, E. (2010): “Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm and non-stochastic simulations,” *Journal of Economic Dynamic and Control*, 34, 36–41.