

Accounting for Idiosyncratic Wage Risk Over the Business Cycle*

Alisdair McKay[†] and Tamás Papp[‡]

March 21, 2012

Abstract

We demonstrate that wage volatility, measured as the cross-sectional variance of wage changes in PSID data, is counter-cyclical. We quantify this relationship by estimating the regression coefficient of wage volatility on the national unemployment rate in a multilevel Bayesian model, then decompose this coefficient into three main factors. During a recession, wage volatility increases substantially among those workers experiencing spells of unemployment: the cyclical changes in the variance within this group explain about 55% of the cyclical variation in wage volatility. The variance within the group not experiencing unemployment explains 18%. Finally, an increase in the fraction of workers experiencing unemployment explains 25%.

We show that a calibrated search-and-matching model of the labor market with on-the-job search gives a good account of the cyclical variation in idiosyncratic wage risk among those experiencing unemployment and of the composition effect over the business cycle. We show that in our model, this result is driven mostly by fluctuations of the reservation wage in response to labor market conditions.

JEL classification: C11, E24, E32, J64

*We thank Christian Haefke, Nobuhiro Kiyotaki, Per Krusell, Monika Merz, Michael Reiter, Christopher Sims, participants in seminars at the Boston University, the Federal Reserve Banks of Richmond and San Francisco, CERGE-EI, CESifo, NORMAC, UCLA, UCSD, and the Vienna Macro Breakfast. Any remaining errors are our own. This research is supported by the Austrian Central Bank Grant No: 14076.

[†]Boston University. amckay@bu.edu

[‡]Institute for Advanced Studies, Vienna. tpapp@ihs.ac.at

1 Introduction

Individuals face a variety of economic risks over their lives. To what degree are these risks shaped by the aggregate economy? And what are the channels that link the aggregate economy with the risks and opportunities that individuals face? In this paper, we empirically explore the extent and sources of business cycle fluctuations in the wage risk faced by individuals. We then consider the theoretical underpinnings of the patterns of counter-cyclical variation in risk that we document.

Counter-cyclical variation in idiosyncratic risk to the components of labor income (employment, hours, wages) has well-known macroeconomic consequences. Recessions are not borne equally by the population and the concentration of economic shocks in a recession tends to reduce the average level of utility and raise the average marginal utility in the economy. These effects make counter-cyclical variation in idiosyncratic risk an important component of the analysis of the potential benefits of aggregate stabilization and asset pricing in models with heterogeneous households facing incomplete insurance markets.¹ In particular, research has shown that household heterogeneity and counter-cyclical variation in idiosyncratic risk can increase the welfare cost of business cycles by orders of magnitude above Lucas’s (1987) original estimate.² To our knowledge, the literature has not investigated the sources of counter-cyclical variation in idiosyncratic risk or the economic mechanisms that drive it. A better understanding of the economic mechanisms at work will allow us to develop structural models of the labor income process that can be used for aggregate stabilization policy experiments that incorporate the link from aggregate labor market conditions to the risks faced by households.

To quantify idiosyncratic risk, we focus on the cross-sectional variance of innovations to log wages in PSID data. Our choice of this cross-sectional moment, which we refer to as “wage volatility,” is motivated by several factors. We focus on wages, as opposed to incomes, because wages are often taken to be the exogenous source of risk in macroeconomic models. We choose to study the dispersion in growth rates, as opposed to levels, because they are a plausible indicator of the size of unanticipated and uninsurable shocks to individuals. Finally, we use the variance to summa-

¹For the potential benefits of aggregate stabilization see Imrohorglu (1989), Atkeson and Phelan (1994), Krusell and Smith (1999), Beaudry and Pages (2001), and references in footnote 2. For asset pricing see Mankiw (1986), Constantinides and Duffie (1996), Storesletten et al. (2007), and Krueger and Lustig (2010).

²See Storesletten et al. (2001); Krebs (2003, 2007); De Santis (2007); Krusell et al. (2009).

alize the dispersion because it allows us to decompose the overall dispersion into different sources. While we find wage volatility to be an important and useful feature of the wage process, there are other features that are also of importance to individuals that we do not consider here, notably the persistence of the shocks. A fully satisfactory structural model of the wage process should generate realistic predictions on these dimensions as well.

We measure the counter-cyclical variation in idiosyncratic wage risk by relating wage volatility in a given year to the national unemployment rate. We do this in the context of a Bayesian multilevel model. The main result of section 2 is the demonstration of a strong comovement between wage volatility and the unemployment rate. Our point estimates imply that wage volatility increases from a level of 0.064 to 0.090 as the unemployment rate rises from three percentage points below its mean to three percentage points above the mean. Our results in section 2 are related to work by Storesletten et al. (2004) who estimate an income process that allows the variance of shocks to differ between expansions and contractions. The key to the Storesletten et al. estimation technique is the recognition that data on incomes from 1967 onwards will contain information about persistent income shocks received since the 1930s. As one of our goals is to understand the sources of the counter-cyclical variation in risk, we cannot use the Storesletten et al. estimation technique because covariates, such as data on unemployment spells, are only available for the years since 1967 when our sample begins. In comparison to Storesletten et al., we use information on a more limited timespan of aggregate fluctuations. Nevertheless, we identify a clear and significant counter-cyclical pattern in wage volatility. While the focus of our paper is on wages, we note in section 2.4 that when we apply our methods to data on incomes the resulting estimates are in line with those of Storesletten et al..

The next step in our analysis is to investigate the sources of counter-cyclical wage volatility in Section 3. One possible explanation is that unemployment shocks are associated with substantial and variable effects on wages as a result of losses of firm-specific human capital or changes in job-specific productivities. As the incidence of unemployment increases in a recession, more individuals are affected by these shocks leading to greater wage volatility in the aggregate. To explore the role that this explanation plays in generating counter-cyclical wage volatility, we partition the sample into two groups according to whether an individual has experienced unemployment in the previous

two years (i.e. the years over which we take the the first-difference in wages). We then perform a variance decomposition exercise that relates the total variance of wage growth rates (i.e. wage volatility) to the within group variances, the within group means and the group sizes and we explore how these components vary with the unemployment rate. Our aim is to decompose the counter-cyclical variation in wage volatility as opposed to decomposing the entire variance of wage volatility across time. To accomplish this, we construct a decomposition of the covariance between wage volatility and the unemployment rate in a way that can be interpreted as decomposing the OLS regression coefficient of wage volatility regressed on the unemployment rate.³

As the incidence of unemployment varies over the cycle, the relative sizes of our two groups varies with more individuals entering the “unemployment” group in a recession. We find that this composition effect generates 25% of the counter-cyclical variation in wage volatility. We find that the majority of the counter-cyclical variation in wage volatility is driven by counter-cyclical movements in wage volatility among those experiencing unemployment spells. This within group variance explains 55% of the total despite the fact that this group makes up only around a quarter of our sample. A further 18% is explained by counter-cyclical movements in the within group variance of those not experiencing unemployment.

Bayesian multilevel models — also known as hierarchical or mixed models — have received little attention from macroeconomists (Sims, 2007) and we believe that we are the first to apply a multilevel Bayesian model to estimate the relationship between aggregate conditions and microeconomic data. This approach has several advantages for our application. First, the multilevel structure incorporates the dependence of cross-sectional moments or parameters on macroeconomic aggregates in a natural manner. Second, the posterior uncertainty surrounding cross-sectional moments is incorporated into our posterior uncertainty about the relationship between these moments and aggregate conditions. Third, multilevel models allow more precise estimation of cross-sectional moments, which is particularly important when we partition the sample so that there are fewer observations per year in each cell of the partition. Finally, posterior sampling allows us to quantify

³Decomposition techniques have been used to understand secular trends in inequality. Notable examples are Juhn et al. (1993) and Lemieux (2006). The sociology literature has also explored trends in income inequality and recently Western and Bloome (2009) have shown how to construct standard errors for Lemieux’s decomposition using Bayesian methods.

the posterior uncertainty surrounding our decomposition results despite the fact that we perform complex non-linear transformations of our parameters in our variance decomposition.

In section 4, we turn our attention to the economic mechanisms generating the data with a focus on search-and-matching frictions. Search-and-matching models provide natural links between aggregate labor market conditions and the experiences of individual workers. Moreover, researchers have begun incorporating search-and-matching frictions into models with heterogeneous, risk-averse workers, which are models that might ultimately be used to assess the welfare consequences of business cycle fluctuations and aggregate stabilization policies with endogenous labor income risk.⁴ We consider a prototypical search-and-matching model with on-the-job search and ask whether the observed variation in the job-offer and separation rates can explain the wage volatility patterns that we document. The model does a good job of explaining the counter-cyclical variation in the variance of wage innovations among those experiencing unemployment and the composition effect, which we found to be the two largest sources of variation in wage volatility. Through these channels, the theoretical model is able to generate 64% of the total effect in the data.

The key mechanism at work in the model is that the job-separation rate increases and the job-offer rate decreases in a recession, both of which lead an unemployed worker to reduce his reservation wage. We explore this mechanism with an analytical approximation to the elasticity of the steady state reservation wage with respect to changes in labor market conditions and use these results to guide our quantitative analysis. A lower reservation wage means there is a larger support of the wage distribution and this leads to more volatile wages as workers are climbing up and falling off a taller wage ladder. In this way, the model generates counter-cyclical wage volatility despite the fact that the variance of offered wages is constant over the business cycle.

⁴There are a number of papers that incorporate risk-averse workers with incomplete insurance markets into search-and-matching models. Of most relevance to our topic are those that also feature aggregate fluctuations. Gomes et al. (2001) model unemployment in the style of Lucas and Prescott (1974). Recently, a number of authors have studied versions of the Diamond-Mortensen-Pissarides search and matching model with risk averse workers and imperfect consumption insurance (Costain and Reiter, 2005, 2007; Shao and Silos, 2007; Rudanko, 2009, 2011; Nakajima, 2010; Bils et al., 2011; Krusell et al., 2010; Jung and Kuester, 2011). Krusell et al. (2011) analyze the roles of search frictions and productivity shocks in explaining individual labor supply decisions over the business cycle.

2 Wage Volatility and Unemployment

In this section we introduce our multilevel modeling approach and estimate a linear relationship between the volatility of wages and the unemployment rate. The data we use are from the Panel Study of Income Dynamics (PSID) covering income in years 1967 to 1992.⁵ We measure wages as the ratio of annual labor income to annual hours worked and deflate to 1967 dollars using the CPI-Research price index. We restrict the sample to male heads who were between the ages of 25 and 60 and worked at least 320 hours per year. Students, business owners, and self-employed individuals are excluded from the analysis. We focus on the first difference of log wages across years so individuals must be present for two consecutive years in order to be included in our sample. Appendix A provides further details about our sample.

2.1 Multilevel Model

Our statistical model consists of three equations. For an individual i in year t , we model the innovation in log wages as

$$dw_{i,t} \sim N(X_{i,t}\theta + \alpha_t, \sigma_t^2) \quad (1)$$

$$\alpha_t \sim N(Z_t\kappa, \varsigma_\alpha^2) \quad (2)$$

$$\sigma_t^2 \sim N(Z_t\eta, \varsigma_\sigma^2), \quad \sigma_t^2 > 0. \quad (3)$$

The first line states that the change in an individual's log wage, dw , is normally distributed. The mean of this distribution depends on the individual's demographic characteristics such as age and education, which are placed in the vector $X_{i,t}$. We assume that the coefficients on these demographic

⁵Two considerations influence our choice of years to include in our sample. First, our object of interest is the first-difference of log annual wages for which we need data on wages in consecutive years. Therefore we cannot make use of PSID data after 1996 (survey year 1997) after which the PSID switches to a biannual frequency. Second, there is a structural break in PSID wage volatility around 1993, which has also been documented by Heathcote et al. (2010). The timing of this break coincides with the switch to a computer-based survey methodology although it is not clear whether this break represents an actual change in the data generating process or if it is an artifact of the methodological change. In our analysis, we have found that our findings survive if we model this break as a level shift in the wage volatility process. We choose, however, to end our sample in 1992 for the sake of simplicity and ease of exposition. Finally, the switch to the new survey methodology began in survey year 1993 and was completed in survey year 1994 so our 1992 data have some elements of the new methodology. The inclusion of 1992 does not exert a strong influence on our results.

characteristics are common across years. In addition, the mean change in wages varies over time with the α_t term. Finally, we allow the variance of the innovation in wages to vary over time as captured by the σ_t^2 term. σ_t^2 is our measure of wage volatility for year t .

The second and third lines show the multiple levels of our model as we impose structure on the parameters of the wage growth distribution and assume that they are drawn from their own, estimated distributions.⁶ The α_t terms are drawn from a normal distribution, the mean of which is linearly related to aggregate variables, Z_t . We use the national unemployment rate as our measure of aggregate conditions and this, along with a constant, makes up the vector Z_t . Similarly, equation (3) relates the variance of wage growth to aggregate conditions. As in equation (2), the mean of this distribution is linearly related to the national unemployment. Our main object of interest is the second element of η , that is the coefficient on the unemployment rate in equation (3), which we call η_{unemp} . As the variance must be positive, we draw from a truncated normal distribution — in practice, however, the mass of the distribution below zero is so small that this truncation is practically irrelevant.

By estimating all of the parameters of the model jointly, the posterior uncertainty about η reflects the uncertainty about σ_t^2 . By contrast, one can imagine a two-stage estimation procedure in which one first estimates equation (1) and then computes σ_t^2 from the residuals of this first-stage regression and uses these estimates as “data” in estimating equation (3). The difficulty with this two-stage approach is that the standard errors in the second-stage regression do not reflect the fact that σ_t^2 is itself an estimate. The multilevel model avoids this difficulty by estimating both stages at once.

Another advantage of the multilevel model is that it provides sharper estimates of σ_t^2 than one would obtain from the first-stage regression. The reason is that the multilevel model is able to pool information across years if the data suggest that $dw_{i,t}$ in those years are generated by a similar process. Consider the meaning of the parameter ς_{σ^2} . If the unemployment rate were the same in years t and s and $\varsigma_{\sigma^2} \approx 0$, then from equation (3) it follows that $\sigma_t^2 \approx \sigma_s^2$, which is to say that the variance of wage changes in years t and s is the same. If these variances are equal, we can estimate them more precisely by pooling the data from years t and s to estimate the single

⁶See Gelman et al. (2004) and Gelman and Hill (2007) for a discussion of multilevel models.

variance that applies to both years. Alternatively, if ς_{σ^2} is very large, the implication is that even if the unemployment rate is the same in years t and s we have no reason to think that σ_t^2 and σ_s^2 should be related. Therefore, we should estimate σ_t^2 and σ_s^2 separately without pooling the data. In between these extremes, the model can partially pool the data across years by down-weighting the data from year s when estimating σ_t^2 .

The amount of pooling that actually occurs in estimating σ_t^2 depends on the value of ς_{σ^2} which is itself estimated jointly with the other parameters of the model. This is possible because the likelihood depends on the parameters of both levels of the model. If the data do not call for pooling, low values for ς_{σ^2} have low likelihood because the data require (relatively) large errors in equation (3) or require that equation (1) be fit with similar variances, which is at odds with the data. Conversely, if the data call for pooling, high values of ς_{σ^2} have low likelihood because the errors in equation (3) are small so the likelihood can be raised by reducing ς_{σ^2} . By following this logic, the multilevel model is able to pool data across years to the extent called for by the data. In section 3.2.1 we discuss exactly how much sharper our estimates are as a result of this partial pooling.

By including the unemployment rate in equation (3), we allow the model to attribute some of the differences in σ_t^2 across years to changes in the unemployment rate. If the unemployment rate explains some of the variation in σ_t^2 , our statistical procedure automatically tightens our estimate of ς_{σ^2} . Since equation (3) is our prior for estimating σ_t^2 , a lower value of ς_{σ^2} implies a sharper prior on σ_t^2 . This prior, which is itself estimated from data across years, is the mechanism through which the model is able to use information from other years to inform the estimate of σ_t^2 . When the prior becomes more precise, it has a larger effect on the estimate of σ_t^2 and more information is pooled across years. So if equation (3) fits better, ς_{σ^2} falls and more information is pooled. In effect, the inclusion of predictors, such as the unemployment rate, in equation (3) allows the model to identify dimensions on which we expect σ_t^2 to differ between years and therefore pool information more effectively.

2.1.1 Prior Distributions

We need to specify prior distributions for θ , κ , η , ς_α and ς_{σ^2} . Since the first three are regression coefficients, it is natural to use the non-informative (reference) priors

$$p(\theta) \propto 1 \tag{4}$$

$$p(\kappa) \propto 1 \tag{5}$$

$$p(\eta) \propto 1 \tag{6}$$

There is sufficient sample size at each level of the model to make the posterior distribution proper.

In the context of an ordinary (single-level) linear regression, the usual choice for a non-informative reference prior for the variance ς^2 of the error term is $p(\varsigma^2) \propto \varsigma^{-2}$. In the context of multilevel models, however, Gelman (2006) demonstrates that this prior places infinite mass near $\varsigma^2 = 0$, resulting in an improper posterior distribution, then suggests that weakly informative priors are used instead, highlighting the advantages of the conditionally conjugate folded-noncentral- t family. Following the recommendation of Gelman (2006) and Polson and Scott (2011), we use a special case of this family, the half-Cauchy priors

$$p(\varsigma_\alpha) \propto (\varsigma_\alpha^2 + s_\alpha^2)^{-1} \tag{7}$$

$$p(\varsigma_{\sigma^2}) \propto (\varsigma_{\sigma^2}^2 + s_{\sigma^2}^2)^{-1}. \tag{8}$$

with $s_\alpha = s_{\sigma^2} = 1$.

2.1.2 Estimation

We use a Gibbs sampler to draw from the posterior distribution of the parameters of the model. We partition the parameter space into blocks corresponding to θ , α , κ , ς_α , σ^2 , η , and ς_α and sample each block in turn. Many of the sampling steps reduce to drawing from an ordinary linear regression or conjugate distributions, otherwise we use the slice sampler of Neal (2003). Appendix B contains details on the estimation methodology and discusses MCMC convergence.

2.2 Results

	mean	5%	25%	50%	75%	95%
θ_{age}	-0.0014	-0.0016	-0.0015	-0.0013	-0.0012	-0.0011
θ_{edu}	0.0028	0.0019	0.0024	0.0028	0.0032	0.0038
κ_{const}	0.0258	0.0202	0.0236	0.0257	0.0279	0.0313
κ_{unemp}	-0.6402	-1.0127	-0.7858	-0.6368	-0.4904	-0.2911
ς_{α}	0.0145	0.0101	0.0123	0.0141	0.0163	0.0201
η_{const}	0.0771	0.0741	0.0758	0.0771	0.0784	0.0802
η_{unemp}	0.4255	0.2257	0.3428	0.4244	0.5077	0.6311
ς_{σ^2}	0.0091	0.0068	0.0079	0.0089	0.0099	0.0119

Table 1: Posterior means and quantiles for model parameters.

Table 1 shows the posterior means and quantiles of the parameters of our model. The posterior mean and median of η_{unemp} are both 0.42 and the 90% error band extends from 0.22 to about 0.62. η_{unemp} can be thought of as the slope of a least squares regression of σ_t^2 on the unemployment rate. As such, a positive coefficient implies a positive co-movement. We also note that the sign of η_{unemp} is almost unambiguously positive: less than 0.2% of the posterior draws are below zero.

To highlight the positive co-movement between these σ_t^2 and the unemployment rate, the right-hand panel of Figure 1 shows a scatter plot of the posterior median of the σ_t^2 against the unemployment rate. The vertical bars in the figure are 90% error bars for σ_t^2 . To show the uncertainty surrounding our estimate of η , the figure also plots posterior draws of the regression line from equation (3). The logic of the multilevel model is that we parameterize the prior distribution on σ_t^2 and equation (3) represents the parameterization of this prior. The difference between the points in the scatter plot and the sample of regression lines is that the regression lines show the variation in the prior on σ_t^2 and the scatter plot shows the estimated values of σ_t^2 , which are a compromise between the data and the prior.

Even though we are primarily interested in η_{unemp} , the posterior distribution of the other parameters is in line with our expectations: κ_{unemp} is estimated to be negative, which implies that real wage growth is counter-cyclical (see the left-hand panel of Figure 1). In the first-level of the model, the demographic effects show that wage growth declines with age and increases with education.

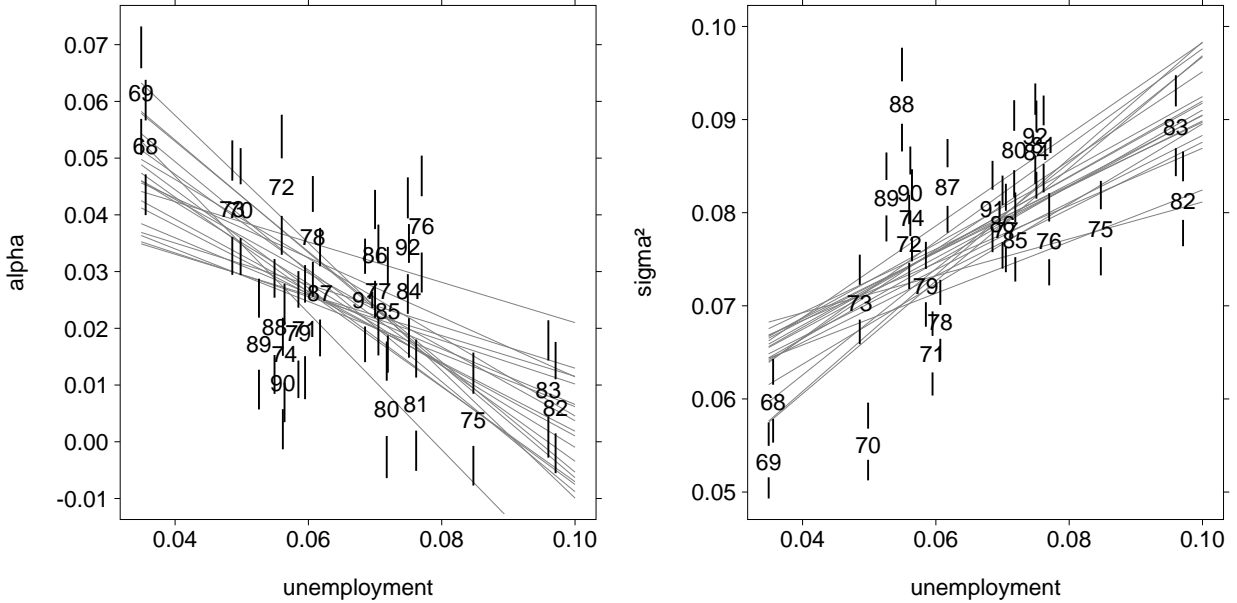


Figure 1: Scatter plot of posterior distribution of α_t (left panel) and σ_t^2 (right panel) versus the unemployment rate in year t . Year labels are placed at the posterior medians, vertical bars extend from the 25% to 5% quantiles and from the 75% to the 95% quantiles. The regression lines are random draws from the posterior distribution.

2.3 High-Frequency Movements in Wage Volatility

Figure 2 shows the time series variation in the unemployment rate and the posterior distribution of σ_t^2 . This figure suggests the possibility that the co-movement between wage volatility and the unemployment rate is driven, at least partially, by slow-moving secular trends. From a theoretical perspective, this low-frequency variation is informative about the sources of wage volatility: changing conditions in the labor market can lead to both higher wage volatility and higher unemployment. Indeed, in section 4 we present a theoretical mechanism that provides a close connection between the unemployment rate and the variance of wage innovations that shows the similar patterns across years.

Nevertheless, we can still investigate whether our estimate of η_{unemp} is a good description of high-frequency movements in wage volatility. To do so, we compare years that are close in time but differ in aggregate conditions, thus minimizing the possibility of slow-moving trends driving result. We use the NBER business cycle dates to identify peaks and then choose the years with

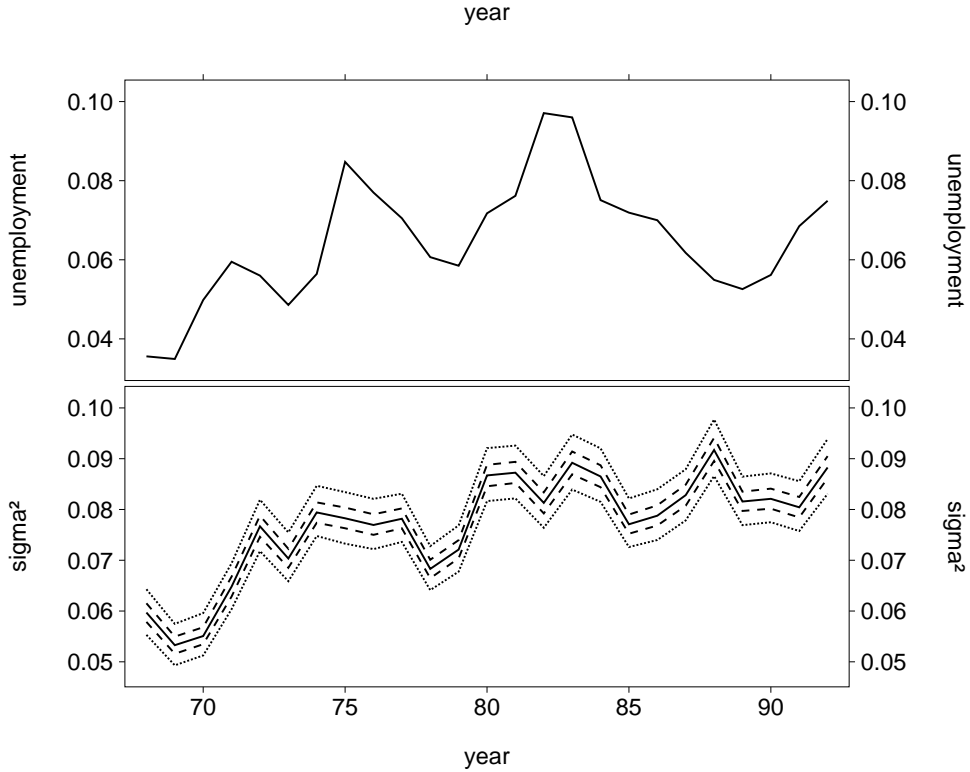


Figure 2: The top panel shows the unemployment rate, the bottom panel shows the posterior quantiles of σ_t^2 . The solid line indicates the median, the dashed lines the 25% and 75% quantiles, and the dotted lines the 5% and 95% quantiles from the posterior.

the lowest and highest unemployment rates between peaks.⁷ The pairs of peaks and troughs we identify are (1969, 1971), (1973, 1975), (1979, 1982) and (1989, 1992). Using posterior draws of σ_t^2 , we calculate the difference in σ_t^2 for each pair and then compute the average of these differences across pairs. For each of the four pairs of years, we compute the ratio of the change in σ_t^2 over the change in the unemployment rate over those years. We then average over the four business cycles. We calculate the posterior distribution of this measure of the slope of the wage-volatility-unemployment relationship. The results in Table 2 show that the mean and median are both 0.31, which is comparable to the estimate of $\eta_{\text{unemp}} = 0.42$ that we obtained from the full sample, suggesting that our results in Section 2.2 are robust.

⁷As we work with annual data, we treat 1980 to 1982 as a single recession.

mean	5%	25%	50%	75%	95%
0.31	0.19	0.26	0.31	0.36	0.43

Table 2: Posterior distribution of $\Delta\sigma_t^2/\Delta u_t$ between paired high- and low-unemployment years. Years are selected based on the highest and lowest unemployment rates between NBER peaks.

2.4 Relation to Previous Work

From the estimated relationship between the unemployment rate and σ_t^2 , we can develop a sense of the changes in wage volatility over the business cycle with the following back of the envelope calculation. Over the course of a typical business cycle in the United States, the unemployment rate fluctuates by roughly 3 percentage points. Our estimate of η_{unemp} then implies that σ_t^2 will increase by 0.013 as the economy moves from the peak of the cycle to the trough.

Many readers will compare our results to the work of Storesletten et al. (2004). Those authors use data on the income of households inclusive of transfers while we use wages of household heads so our results are not directly comparable to theirs. Bearing these differences in mind, one might still ask if our results are plausible in comparison to theirs. Their results imply that the variance of the first-difference in household earnings increases by about 0.041 as the economy moves from expansion to contraction.⁸

To facilitate the comparison of our results to theirs, we also estimate the model using labor income as data instead of wages. All other features of our analysis remain the same. In this case, the posterior mean of η_{unemp} is 1.14 with a 90% error band extending from 0.76 to 1.54. Performing the same back of the envelope calculation as above produces a difference in income volatility of 0.034 as the unemployment rate increases by 3 percentage points. Given the important differences between their methodology and data and ours, we do not find it surprising that we obtain somewhat different results.⁹

⁸Appendix C explains how we reach this conclusion from their results.

⁹The difference in methodology is that they look at the cross-sectional dispersion in earnings across cohorts who have lived through different macroeconomic experiences while we look at the dispersion of the *first-differences* in booms and recessions. The difference in data is that they look at total household income inclusive of transfers, while we look at just the labor income of male heads.

2.5 Impact of Measurement Error

The PSID data on wages are surely affected by measurement error so it is important to consider how our results are affected. Suppose we measure $dw_{i,t} = \widehat{dw}_{i,t} + \varepsilon_{i,t}$ where \widehat{dw} is the true wage growth and ε is measurement error distributed i.i.d. normal with some mean and variance, v_ε . Then it follows from equation (1) that $\sigma_t^2 = \hat{\sigma}_t^2 + v_\varepsilon$, where $\hat{\sigma}_t^2$ is the variance of $\widehat{dw}_{i,t}$. So measurement error of this type will shift the intercept in equation (3), but η_{unemp} is not affected. More generally, if the extent of measurement error varies over time it will affect our result to the extent that it covaries with the unemployment rate.

3 Decomposition

We now investigate the forces that drive the positive co-movement between wage volatility and the unemployment rate. In section 3.1 we lay out a methodology for decomposing the slope of a linear regression of wage volatility on the unemployment rate. The first key step in this decomposition is that we use observed covariates to partition the sample into groups within each year. We can then decompose the total variance of wage growth in a year into variances within groups and the variance between groups. Movements in the total variance over years are then driven by movements in these within- and between-group variances as well as shifts in the group sizes. The second key step is to model (statistically) the mean and variance of wage growth within each cell of the partition in each year. Doing so allows us to capture the posterior uncertainty about the components of the decomposition and therefore assess the uncertainty about the decomposition results themselves.

3.1 Methods

Suppose that we can partition individuals into J groups within each year based on observed covariates. When we apply this methodology below, we form two groups according to whether an individual has experienced any unemployment spells in the preceding two years and so $J = 2$, but we choose to keep our discussion in general terms to emphasize the fact that this methodology could be applied to any partition of the data. In the sums below, the index j always runs from 1 to J .

3.1.1 Decomposing η_{unemp}

We now explain our method for decomposing the slope of a regression of wage volatility on the unemployment rate, which is closely related to the parameter η_{unemp} that we estimated in section 2. Given our partition of the data, suppose we know the fraction of individuals in each group and the mean and variance for wage growth in each group. We can then construct the total variance of wage growth across all individuals as

$$\sigma_t^2 = \sum_j \pi_{t,j} \sigma_{t,j}^2 + \sum_j \pi_{t,j} (\alpha_{t,j} - \alpha_t)^2 \equiv W_t + B_t \quad (9)$$

where $\sigma_{t,j}^2$ is the within-group variance and $\alpha_{t,j}$ is the within-group mean for group j at t . Also, $\pi_{t,j}$ is the fraction of observations in group j at time t , and α_t and σ_t^2 are the mean and variance for all observations.

As is customary, we call the two sums in equation (9) *within-* and *between-group* variances. There are two forces that can raise the contribution of the within-group variance. First, the group fractions, $\pi_{t,j}$ might shift towards groups with higher within-group variances. Second, the within-group variances might increase. In order to separate out these two effects, we further decompose the within-group variance term using the cyclical deviations from a time-series mean. That is to say, let

$$\bar{\pi}_j = \frac{1}{T} \sum_{t=1}^T \pi_{t,j} \quad \text{and} \quad d\pi_{t,j} = \pi_{t,j} - \bar{\pi}_j$$

denote the *average* share of observations in each group and the *deviations* from this average, and similarly

$$\bar{\sigma}_j^2 = \frac{1}{T} \sum_{t=1}^T \sigma_{t,j}^2 \quad \text{and} \quad d\sigma_{t,j}^2 = \sigma_{t,j}^2 - \bar{\sigma}_j^2$$

Then we can write W_t as

$$W_t = \sum_j \bar{\pi}_j \bar{\sigma}_j^2 + \sum_j \bar{\pi}_j d\sigma_{t,j}^2 + \sum_j \bar{\sigma}_j^2 d\pi_{t,j} + \underbrace{\sum_j d\pi_{t,j} d\sigma_{t,j}^2}_{e_t} \quad (10)$$

Equation (10) is exact, but we can think of the last term as the second-order error term of a linear

approximation, so we will denote it by e_t .

Our goal is to decompose the connection between our cyclical indicator, the unemployment rate, and the overall variance of the wage innovations. We find it convenient to use the OLS regression slope for this purpose as a summary statistic, which we define as

$$\hat{\eta}_{\text{unemp}} = \sigma_t^2 // u_t \quad (11)$$

where $//$ is the univariate regression slope, defined for a time series y_t as

$$y_t // u_t = \frac{\sum_t (y_t - \bar{y})(u_t - \bar{u})}{\sum_t (u_t - \bar{u})^2} = \frac{\text{Cov}(y, u)}{\text{Var}(u)}. \quad (12)$$

This is simply the point estimate of an OLS regression of y_t on u_t .¹⁰ Note that $y // u$ is additive in y ,¹¹ and thus it can be applied to both sides of (9) and (10) to obtain an additive decomposition

$$\sigma_t^2 // u_t = \underbrace{\sum_j \bar{\pi}_j (d\sigma_{t,j}^2 // u_t)}_{\text{cyclical variance of groups}} + \underbrace{\sum_j \bar{\sigma}_{t,j}^2 (d\pi_{t,j} // u_t)}_{\text{composition}} + \underbrace{e_t // u_t}_{\text{error}} + \underbrace{B_t // u_t}_{\text{between}} \quad (13)$$

The first term summarizes the effect of the *variance changing within each group* j , this is of course weighted by the proportion of observations within each group. The second term stands for the effect of the number of agents changing within each group: even if we held variances constant, the change in proportions would result in a *composition effect*. Finally, we have the error and between group variance terms. We also find it useful to summarize (13) normalized by its left hand side. This provides a very condensed summary of our results that allow the reader to gauge the contribution of each effect.

Our decomposition could be applied directly to sample moments computed from the data. Using sample moments, however, does not give any sense of the uncertainty surrounding the results. To the

¹⁰The parameter $\hat{\eta}_{\text{unemp}}$ is closely related to η_{unemp} , which we estimated above. In the terminology of multilevel models, the latter is a *superpopulation* parameter, while the former is called its *finite population* counterpart. As we use a flat prior for η_{unemp} , the two are very close numerically. The advantage of using $\hat{\eta}_{\text{unemp}}$ is that it is easy to visualize and can be decomposed additively as we show here. For a discussion of finite vs superpopulation moments, see Gelman and Hill (2007, Chapter 21.2) and Gelman (2005).

¹¹This is easy to check from (12).

extent that we partition the sample into small groups, our uncertainty about the sample moments can be substantial so it is important to account for this uncertainty. We solve this problem by extending our statistical model to estimate the within-group means and variances. We then draw these means and variances from the posterior distribution and use these draws to compute the terms in equation (13). In doing so, we are calculating the posterior distribution of a function of our parameters.

3.1.2 Extending the Model

We now extend our statistical model to the mean and variance of wage growth within each cell of the partition. For an individual i who is in cell j of the partition in year t , the extended model specifies the following distribution for wage growth

$$dw_{i,t} \sim N\left(X_{i,t}\theta + \alpha_{t,j[i,t]}, \sigma_{t,j[i,t]}^2\right) \quad (14)$$

$$\alpha_{t,j} \sim N(Z_t k_j, \varsigma_{\alpha j}) \quad \forall j \quad (15)$$

$$\sigma_{t,j}^2 \sim N(Z_t \eta_j, \varsigma_{\sigma^2 j}), \quad \sigma_{t,j}^2 > 0 \quad \forall j. \quad (16)$$

Equation (14) states that the change in an individual's log wage, $dw_{i,t}$ is normally distributed. As before, the mean of this distribution depends on the individual's demographic characteristics and the coefficients on these demographic characteristics are assumed to be common across groups and across years. In addition, the mean change in wages varies over time and across groups with the $\alpha_{t,j[i,t]}$ term. The notation $j[i,t]$ refers to the group index for the group that individual i is a member of at time t . We also allow the variance of the innovation in wages to vary over time and across groups as captured by the $\sigma_{t,j[i,t]}^2$ term. Equations (15) and (16) model the within group means and variances, respectively. Again, these components are related to aggregate conditions captured by Z_t . Priors are independent for each group, and otherwise have the same form as in section 2.1.1.

3.2 Decomposing by Unemployment Experience

We apply our decomposition method to the PSID data partitioned by unemployment experience. In particular for an individual i at time t , $dw_{i,t}$ refers to wage growth between years $t - 1$ and t . We assign an individual to the “no unemployment” group if this individual reports zero hours of unemployment for both years $t - 1$ and t . Those reporting positive hours of unemployment in year $t - 1$ or year t are assigned to the “unemployment” group.¹²

3.2.1 Decomposition Results

Table 3 shows the posterior distribution of the decomposition (13), in levels and in normalized form. Panel a) shows the contribution of each component to the total slope $\hat{\eta}_{\text{unemp}}$ and Panel b) expresses these contributions in percentage terms. From Panel b), one can see that counter-cyclical fluctuations in the within group variances contribute 73% of the total with the unemployment group contributing the bulk of this (55%). Most of the remainder comes from the composition effect that arises because the unemployment group has a larger variance at all times and the size of this group increases during a recession. This composition effect contributes 25% of the total. Finally, the error in the decomposition and the between variance contribute next to nothing. The table also shows quantiles for these fractions and one can see that the posterior uncertainty does not change the overall message.

We view the results in Table 3 as our main empirical results and we now present additional findings from the extended model to explain the forces that drive our results. Figure 3 plots our estimates of $\sigma_{t,j}^2$ for both groups. From these plots, one can see that the comovement with the unemployment rate is much stronger in the unemployment group. While the fluctuation of variance is relatively small in the no unemployment group, this group is about four times the size of the unemployment group and thus its contribution is scaled up by a factor of about 0.8 when the term $d\sigma_{t,j}^2$ is multiplied by $\bar{\pi}_j$ in equation (13) as opposed to 0.2 for the unemployment group.

One way of understanding the benefit of using the multilevel model is to compare the posterior uncertainty of the $\sigma_{t,j}^2$'s to the posterior uncertainty that would result without any pooling (e.g.

¹²See Appendix A for more details.

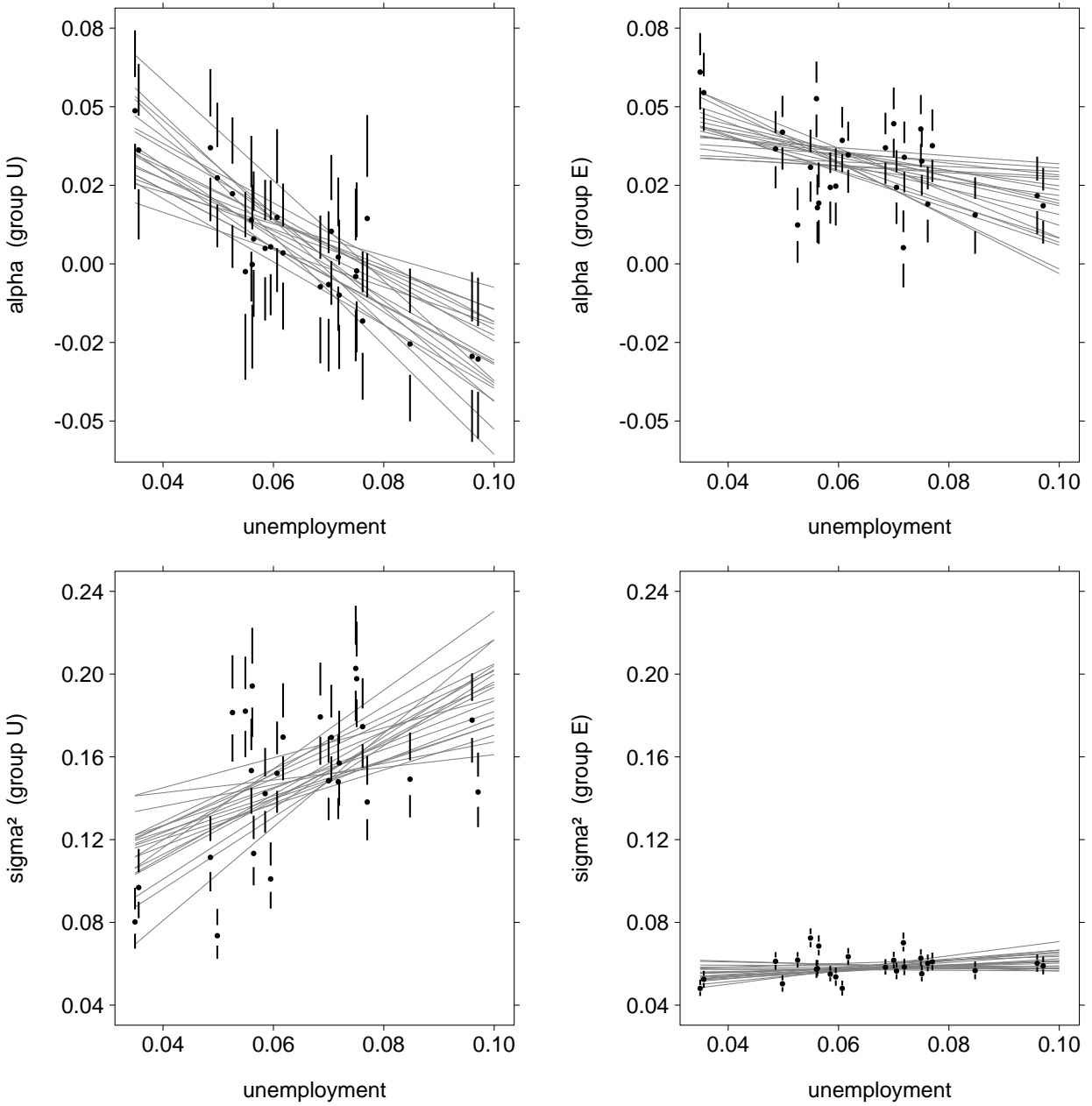


Figure 3: Scatter plot of posterior distribution of α_t (top panels) and σ^2 (bottom panels) versus the unemployment rate in year t , for the “unemployment” (U, left panels) and “no unemployment” (E, right panels) groups. Dots are placed at the posterior medians, vertical bars extend from the 25% to 5% quantiles and from the 75% to the 95% quantiles. The regression lines are random draws from the posterior distribution.

	mean	5%	25%	50%	75%	95%
$d\sigma_{\text{U}}^2$	0.2361	0.1836	0.2147	0.2361	0.2587	0.2900
$d\sigma_{\text{E}}^2$	0.0775	0.0308	0.0574	0.0772	0.0971	0.1244
composition	0.1040	0.0987	0.1018	0.1039	0.1060	0.1095
between	0.0074	0.0014	0.0043	0.0070	0.0100	0.0147
error	0.0018	-0.0075	-0.0019	0.0016	0.0057	0.0113

a) decomposition of regression slope

	mean	5%	25%	50%	75%	95%
$d\sigma_{\text{U}}^2$	0.5533	0.4732	0.5206	0.5539	0.5866	0.6320
$d\sigma_{\text{E}}^2$	0.1791	0.0810	0.1404	0.1809	0.2193	0.2722
composition	0.2464	0.2060	0.2269	0.2433	0.2626	0.2952
between	0.0174	0.0033	0.0103	0.0165	0.0236	0.0352
error	0.0038	-0.0186	-0.0045	0.0039	0.0131	0.0248

b) normalized decomposition

Table 3: Absolute (top) and normalized (bottom) decomposition of the regression of wage volatility on the unemployment rate. The rows of the both tables are the cyclical variation of the volatility of the unemployed ($d\sigma_{\text{U}}^2$) and employed ($d\sigma_{\text{E}}^2$) groups, followed by the composition effect, and the within and between variance, as shown in equation (13). Note that even though the “unemployed” (U) group makes up only $\approx 20\%$ of the sample, it is responsible for more than half of the cyclicity of the variance. In contrast, the “no unemployment” (E) group only explains about $1/5$ of the cyclicity. Finally, the composition effect explains the rest (about $1/4$), and the other terms have a negligible effect.

from a model with a single level — we specify the details of this model in Appendix B.3). This comparison shows how much posterior uncertainty is eliminated by partially pooling information across years. Figure 4 displays this comparison, quantifying uncertainty using posterior variances, for each year. For both groups, the multilevel model tightens our estimates of $\sigma_{t,j}^2$. The fact that there is a larger benefit of pooling in the unemployment group reflects the fact that there are fewer observations in this group and so information contained in the hierarchical prior has a larger impact on the posterior.

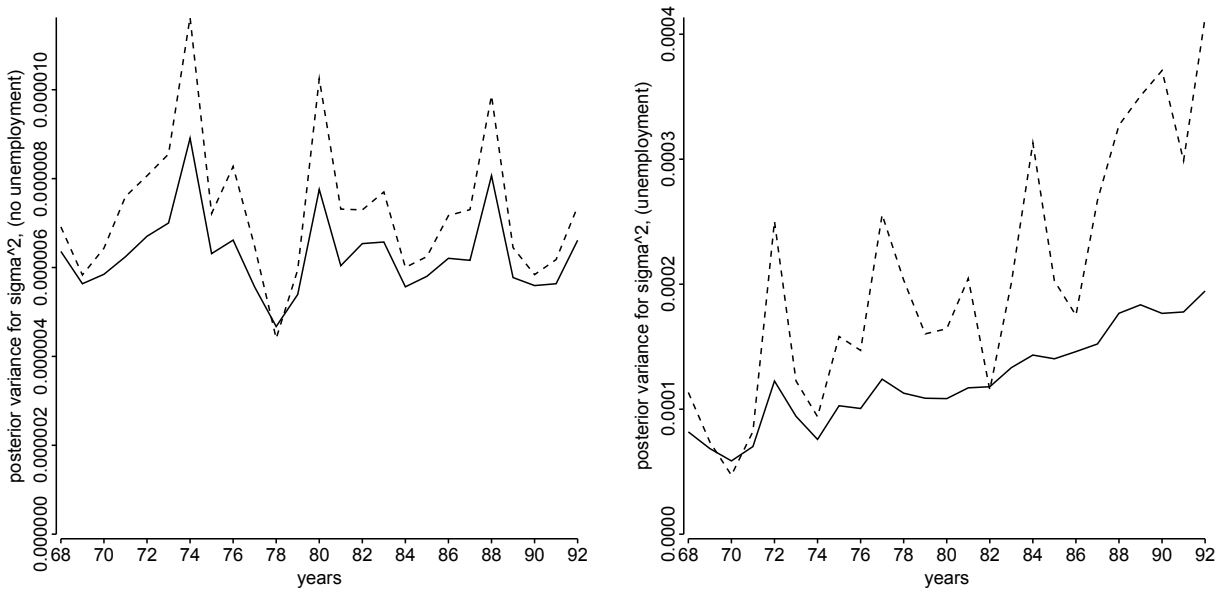


Figure 4: Posterior variance the $\sigma_{t,j}^2$ parameters (by year) for the no unemployment group (left panel) and unemployment group (right panel), in the multilevel (solid lines) and non-pooling models (dashed lines). The reduction in posterior variance demonstrates that the parameters of interest are estimated with a higher precision in the multilevel model.

4 Modeling Wage Risk Over the Business Cycle

We now demonstrate that a standard search-and-matching model of the labor market can generate most of the counter-cyclical idiosyncratic wage risk that we have documented as a result of business cycle variation in job-offer and job-separation rates. We focus on a partial equilibrium analysis in which we take the job-offer and separation rates and the distribution of offered wages as given. In doing so, we sidestep an important, unresolved question about the magnitude of fluctuations in labor market conditions (see Shimer, 2005; Costain and Reiter, 2008).

Whether the model can generate the counter-cyclical variation in wage volatility that we have observed in the data is a quantitative question, which we explore with a calibrated version of the model. Our analysis shows that business cycle fluctuations in the rates at which workers contact firms and separate from existing matches induces endogenous movements in the reservation wage that leads to counter-cyclical wage volatility among those workers experiencing unemployment. The model is also able to generate a composition effect. As movements in the reservation wage

are an important source of changes in wage volatility, we explore the sensitivity of the steady state reservation wage to changes in labor market conditions as a way of understanding the factors that determine the strength of this mechanism. This analysis guides our calibration of the model for the quantitative analysis that follows.

4.1 Wage-Ladder Model

Consider a partial equilibrium wage ladder model, where the log wage for individual i at time t is the sum of an *aggregate* component p_t , an *individual-specific* component $a_{i,t}$, and a *match-specific* component $x_{i,t}$

$$w_{i,t} = p_t + a_{i,t} + x_{i,t}.$$

A job offer is a draw of a match-specific component from a wage-offer distribution $F(\cdot)$, which is exogenous and time invariant. Workers have logarithmic utility, and discount factor β . Unemployed workers receive a (log) unemployment benefit $p_t + a_{i,t} + b$. An unemployed worker receives a job offer with probability λ_t at time t . An employed worker receives a job offer with probability $\gamma\lambda_t$ and moves to unemployment with probability δ_t . γ is a parameter that controls the relative efficiency of on the job search and setting $\gamma = 0$ gives us a variant of the McCall (1970) model. λ_t and δ_t are the only exogenous sources of variation at the aggregate level.

A worker's decision problem is whether to accept a job offer. Employed workers accept any wage that is higher than their current wage. Unemployed workers suffer a reduction in search efficiency when they accept a job offer and therefore adopt a reservation wage strategy. For ease of exposition, we first present the steady state version of the model in which λ and δ are constant. The Bellman equations for the unemployed and employed workers are

$$U(p, a) = p + a + b + \beta \mathbb{E}_{(p', a', x')} \left[\lambda \max(W(p', a', x'), U(p', a')) + (1 - \lambda)U(p', a') \right] \quad (17)$$

$$W(p, a, x) = p + a + x + \beta \mathbb{E}_{(p', a', x')} \left[\gamma \lambda \max[W(p', a', x'), W(p', a', x)] + \delta U(p', a') + (1 - \gamma\lambda - \delta)W(p', a', x) \right] \quad (18)$$

In Appendix D.1 we show that the value functions separate linearly into a component that depends on $x_{i,t}$ and one that depends on p_t and $a_{i,t}$, and thus the worker's search strategy can be analyzed independently of p_t and $a_{i,t}$.

4.2 Steady State Comparative Statics for the Reservation Wage

The response of the reservation wage to the job-offer and separation rates turns out to be the most important source of variation in wage volatility in this model (see Figure 6 below). Here, we explore the sensitivity of the steady state reservation wage to variation in λ and δ to develop intuition for the features of the model that govern the magnitude of the variation in the reservation wage. This analysis also informs our choice of calibration moments for the full dynamic model in the following subsection. Studying the steady state reservation wage is a useful guide to the behavior of the reservation wage in the full dynamic model and this can be verified *ex post*.

Unemployed workers use a reservation rule, accepting jobs when $x \geq x^*$, where

$$x^* = b + \beta(1 - \gamma)\lambda \int_{x^*}^{\infty} \frac{\bar{F}(x)}{1 - \beta + \beta\delta + \beta\gamma\lambda\bar{F}(x)} dx \quad (19)$$

defines x^* and $\bar{F}(x) = 1 - F(x)$ is the tail probability (see Appendix D.1). With some slight abuse of terminology, we refer to x^* as the reservation wage. Introducing

$$r = \frac{1}{\beta} - 1 \quad \text{and} \quad \chi = \frac{\lambda}{r + \delta}$$

and integration by parts allows us to rewrite (19) as

$$x^* = b + (1 - \gamma)\chi \int_{x^*}^{\infty} \frac{x - x^*}{(1 + \gamma\chi\bar{F}(x))^2} F'(x) dx \quad (20)$$

From (20) it is clear that when the wage offer distribution, the unemployment benefit and the relative search intensity are fixed, the effect of all other labor market conditions on the reservation wage can be summarized in χ . This is a useful result because the variation in λ and δ can be summarized by their effect on χ .

Lemma 1. *The steady state elasticity¹³ of the reservation wage x^* with respect to χ , defined in equation (20), can be calculated up to a first-order approximation as*

$$\frac{\hat{d}x^*}{\hat{d}\chi} = \underbrace{\frac{\chi^*}{1 + \chi^*}}_A \underbrace{\left(\frac{\bar{x}}{x^*} - 1\right)}_B \underbrace{(1 - \gamma(1 + \chi^* M_\gamma(F, x^*)))}_C \quad (21)$$

where

$$\chi^* = \chi \bar{F}(x^*) = \frac{\lambda \bar{F}(x^*)}{r + \delta} = \frac{\lambda_{UE}}{r + \delta}, \quad (22)$$

and λ_{UE} is the observed unemployment-employment transition probability,

$$\bar{x} = \mathbb{E}_x[x \mid x \geq x^*]$$

is the expected value of x , conditional on the worker accepting the offer, and

$$M_\gamma(F, x^*) = \frac{2M_2(F, x^*)}{M_1(F, x^*)} - 1, \quad (\text{which is in the interval } [0, 1] \text{ for all distributions})$$

characterizes the tail shape of the distribution, with

$$M_i(F, x^*) = \frac{\int_{x^*}^{\infty} (\bar{F}(x))^i dx}{\bar{F}^i(x^*)} = \int_{x^*}^{\infty} (\bar{F}^*(x))^i dx \quad \text{where} \quad \bar{F}^*(x) = \frac{\bar{F}(x)}{\bar{F}(x^*)}.$$

Proof. See Appendix D.2. □

Equation (21) informs our calibration strategy for our quantitative analysis of the model. Since $r \ll \delta$,

$$\chi^* \approx \frac{\lambda_{UE}}{\delta} = \frac{1 - u}{u},$$

the first term, labeled A, is very close to the steady state unemployment rate. Also, $\chi^* = \lambda_{UE}/(\delta+r)$ can be calibrated directly from observed worker flows and the interest rate.

¹³In this paper, we use the notation

$$\hat{d}\chi = d \log \chi$$

for proportional change.

The second term in equation (21) is

$$B = \frac{M_1(F, x^*)}{x^*} = \frac{\bar{x}}{x^*} - 1 \quad (23)$$

which measures the expected gain to an unemployed worker of receiving an offer relative to the reservation wage. The value of this term depends on the distribution F , but it can be calibrated *independently of the job finding and separation probabilities*. Hornstein et al. (2012) show that in a wide class of search models, the ratio of the average accepted wage and the reservation wage (the so-called *mean-min ratio*) can be calibrated using only labor market flows and the replacement ratio (the ratio of the unemployment benefit and the average wage), and that bringing the mean-min ratio in line with the data requires a low unemployment benefit. The equation above features a *log* mean-min ratio, but it has a similar interpretation. From this term, we take away the implication that the reservation wage will not be sensitive to χ if the mean-min ratio is too low.

Finally, the third term in (21) is

$$C = 1 - \gamma (1 + \chi^* M_\gamma(F, x^*)).$$

This captures the effect of on-the-job search: more effective on-the-job search, a larger value of γ , reduces the elasticity $\hat{d}x^*/\hat{d}\chi$. One implication is that it is important to include on-the-job search in the model so as not to overstate the elasticity of the reservation wage. A second implication is how the shape of the offer distribution affects the response of the reservation wage to labor market conditions. The term M_γ ranges from 0 to 1, and captures how self-similar the shape of the distribution is: the two extremes are the exponential distribution, where

$$F(x^*) = \exp(-ax^*) \quad M_i = a^{-i} \quad M_\gamma = 0$$

for some parameter $a > 0$, and the distribution with a single mass point at some $x_1 > x^*$, where

$$F(x^*) = 1(x \leq x_1) \quad M_i = x_1 - x^* \quad M_\gamma = 1$$

In this paper we focus on the normal distribution, which is in between these two extremes. Assessing the effect of M_γ is made easier by the fact that it is *invariant to scale and shift transformations*, and thus it can be discussed independently of the mean and variance of the distribution of F . In this case, M_γ depends only on the fraction of offers that are rejected.¹⁴ Figure 5 shows how M_γ depends on the fraction of offers that are rejected. In our calibration strategy, we target this rejection rate.

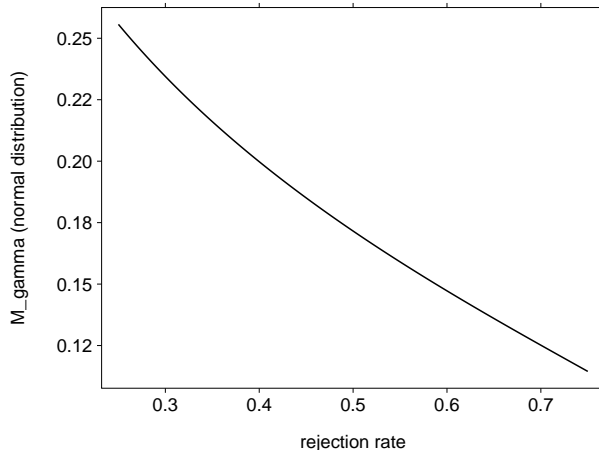


Figure 5: M_γ for the standard normal distribution, as a function of the fraction of offers rejected. See Appendix D.3 for a derivation of this function.

4.3 Calibration

We now calibrate the wage ladder model in order to quantitatively explore its predictions for wage volatility over the business cycle. Our time period is 1/48 of a year. Table 4 summarizes the calibration. We start with parameters that can be matched to observables in a straightforward manner: we calibrate r to a 4% annual interest rate, set the separation rates δ to match the average of those in Shimer (2007) over the years 1968 to 1992, and use a result of Nagypál (2005) to calibrate the relative search efficiency γ in a manner that is independent of the distributions,

¹⁴This is in fact simple to show: let

$$z = \frac{x - \mu}{s}$$

for some fixed μ and s , and F_z be the distribution transformed from F accordingly, similarly for z^* from z . Since

$$M_i(F, x^*) = s M_i(F_z, z^*) \quad \text{for } i = 1, 2,$$

M_γ and thus C are unaffected, and only depend on z^* .

	description	value	calibrated to
r	interest rate	$(1.04)^{1/48} - 1$	real interest rate
δ	separation rate	9.78×10^{-3}	separation rate in Shimer (2007)
γ	relative OTJ search efficiency	0.151	steady state flows, see Section E.2
ρ	autoregressive coef. for $a_{i,t}$	0.997	moments of continuously employed, see Section E.1
σ_a	std. dev. for $a_{i,t}$ innovation	0.000385	
σ_ϵ	std. dev. of measurement error	0.148	
λ	offer arrival rate	0.245	job finding rate in Shimer (2007)
s	std. dev. of offer distribution	0.337	matching level of variance
b	log unemployment benefit factor	-1.619	determined by steady state (20)

Table 4: Calibrated parameters. Time unit is 1/48 of a year.

using only the transition probabilities, particularly the average monthly job-to-job transition rate of 2.6% as reported by Fallick and Fleischman (2004). The details are discussed in Appendix E.2.

Our model has two sources of wage changes:¹⁵ those coming from worker-specific differences captured by the process $a_{i,t}$ and those coming from the search and matching structure, which are reflected in values of $x_{i,t}$. We also allow for some of the variance in wage growth rates that we observe in the data to be driven by measurement error. We assume that $a_{i,t}$ follows an AR(1) process, the parameters of which are stable throughout time; we also assume that the extent of measurement error is also stable throughout time. As these components are not time-varying, their main role is to create variation in wage growth rates that is not related to search and matching. Considering that in our model $x_{i,t}$ does not change for workers who are continuously employed on the same job, focusing on this subset of the data allows us to calibrate an AR(1) process for $a_{i,t}$ using variances and autocorrelations. We discuss the calibration of these processes in Appendix E.1.

We assume that the distribution F is normal and normalize the mean to zero. We choose the variance, s^2 , so that the model matches the unconditional variance of wage growth rates in our PSID sample.

We make the model match the job finding rate $\lambda_{UE} = \lambda \bar{F}(x^*)$ using the job finding probabilities

¹⁵The aggregate log wage p_t also changes, but, as shown in Section D.1, does not affect the reservation wage. As all wages are affected equally by p_t , it will not affect the dispersion of wage growth rates and so we do not need to specify a process for it.

reported by Shimer (2007).¹⁶ In order to pin down the levels of both λ and $\bar{F}(x^*)$ we need an additional moment. Our analysis in section 4.2 shows that the offer acceptance probability $\bar{F}(x^*)$ is an important component of the model so we directly calibrate this feature of the model from survey evidence. Blau and Robins (1990) use data from the Employment Opportunity Pilot Project baseline survey, which reflected job-search experiences in 1979 and 1980 and found that unemployed individuals received an average of 0.18 offers per week and 10 percent of the unemployed individuals accepted an offer of employment per week. These findings imply that around 5/9 of offers received by unemployed individuals are accepted. The acceptance rate could be somewhat higher if single individuals received multiple job offers within a week.¹⁷ Krueger and Mueller (2011) conducted a survey of Unemployment Insurance recipients in the state of New Jersey in late 2009 and found that about 60% of job offers are accepted. As the labor market was particularly slack in 2009, we would expect the steady state acceptance rate to be somewhat lower than 60%. Therefore, we target $\bar{F}(x^*) = 5/9$.

Finally, we note that the log unemployment benefit factor b is determined by the above parameters via (20). As is emphasized by Hornstein et al. (2012), the model cannot deliver a reasonable mean-min ratio for wages if the unemployment benefit is calibrated in the standard way. The benefit replacement rate in our calibration is 13%, quite a bit lower than typical calibration strategies. A low value of b is needed to generate dispersion in wages and wage growth rates: with a dispersed offer distribution, the returns to search are large and in order to prevent unemployed workers from rejecting a large fraction of offers we must make the cost of search large as well, hence a low value of b . Our analysis of the steady state elasticity of the reservation wage shows that the unemployment benefit does not affect the sensitivity of the reservation wage except through the mean-min ratio. Therefore, our calibration strategy focuses on matching a measure of wage dispersion rather than

¹⁶The Shimer (2007) and Fallick and Fleischman (2004) data are monthly transition probabilities and we must convert them to weekly probabilities. Suppose L is a monthly probability, then $\ell = -\log(1 - L)$ is the continuous-time arrival rate. $1 - \exp(-\ell/4)$ is the weekly transition probability implied by the continuous-time process, which is what we use in our calibration. For the Shimer data, we take the average of this value over the years 1968 through 1992, which corresponds to our PSID sample period.

¹⁷If offers arrive as a Poisson process within the week then the probability of not accepting an offer during the week is

$$1.0 - 0.1 = e^{-0.18} + \frac{e^{-0.18}0.18}{1!}F(x^*) + \frac{e^{-0.18}0.18^2}{2!}[F(x^*)]^2 + \dots$$

where 0.18 is the weekly offer arrival rate and $[F(x^*)]^k$ is the probability of rejecting k offers. Solving this equation for $F(x^*)$ yields $F(x^*) = 0.415$, which is only slightly below 4/9.

unemployment benefits.

4.4 Wage Volatility Over the Business Cycle

We conduct the following experiment to assess the ability of the model to generate wage volatility patterns like those we have documented in the data. We allow the job-offer and separation rates, λ and δ respectively, to vary over time so that the model reproduces the observed job-finding and separation probabilities from 1948 to 2007 reported by Shimer (2007). When we do this, we assume that all other parameters of the model are constant through time. Changes in the transition rates will, however, affect the worker's choice of reservation wage. In solving the worker's problem, we assume that the worker has perfect foresight for the paths of λ_t and δ_t . After 2007, we assume that the parameters return to their steady state values. As we are focusing on the wage dynamics from 1968 to 1992, this return to steady state occurs 15 years after the end of the period we are interested in. We can then simulate a population of workers from 1948 to 1992. We use the steady state theoretical wage distribution to initialize the simulation in 1948 and the period 1948-1967 acts as burn-in period for our simulation results. We then compare the wage dynamics from 1968 to 1992 to those we found in the PSID data in Sections 2 and 3. Changes in λ_t and δ_t can affect wage dynamics directly as workers move up and fall off the wage ladder more or less quickly, and indirectly through the reservation wage, as discussed in section 4.2.

The job-finding rate is the product of the job-offer rate and the probability that an unemployed worker accepts an offer, $\bar{F}(x_t^*)$. Pro-cyclical movements in the reservation wage imply that the job-offer rate must be more volatile than the observed job-finding rate. At first glance, this would seem to make the unemployment volatility puzzle (Shimer, 2005) more severe, however, Menzio and Shi (2011) have shown that on-the-job search can lead to more volatility in unemployment albeit in a richer model with directed search. With regard to the separation rate, we assume that workers cannot quit into unemployment when we solve for the reservation wage and simulate the model. Therefore, all separations are exogenous and the separation rate that is fed into the model is the same that we observe in the data. In our model, a worker would want to quit into unemployment if he held a low-wage job and then labor market conditions improved leading the reservation wage to

	mean	5%	25%	50%	75%	95%	simulation
$d\sigma_U^2$	0.2361	0.1836	0.2147	0.2361	0.2587	0.2900	0.1913
$d\sigma_E^2$	0.0775	0.0308	0.0574	0.0772	0.0971	0.1244	-0.0212
composition	0.1040	0.0987	0.1018	0.1039	0.1060	0.1095	0.0798
between	0.0074	0.0014	0.0043	0.0070	0.0100	0.0147	0.0227
error	0.0018	-0.0075	-0.0019	0.0016	0.0057	0.0113	-0.0010

Table 5: Estimated and simulated decomposition. The table contains the empirical decomposition from Section 3.2.1 (top panel of Table 3), extended with the simulated decomposition.

rise above his current wage. Allowing workers to quit into unemployment results in little bursts of separations when the reservation wage increases. To offset these endogenous separations, we would need to reduce the rate of exogenous separations at this time, but there is only so much room to do so since the separation rate is already low. As a result, we were unable to match the observed separation rates exactly when we allow these quits to occur.

We find that the model can generate a substantial fraction of the counter-cyclical volatility in wages, which can be seen in Figure 6. The slope of a linear regression relating the variance of wage growth rates to the unemployment rate is 0.271 in simulated data, while in our empirical estimates of η_{unemp} in section 2 we found a posterior mean of 0.426. Table 5 shows a decomposition of this relationship between those experiencing unemployment and those not experiencing unemployment as we did empirically in section 3. We see here that model generates cyclical variation in wage volatility from the variance within the unemployment group and the composition effect and the magnitudes of these effects are close to our empirical estimates. The contribution from the variance among those not experiencing unemployment, however, is actually slightly negative as opposed to mildly positive in the empirical estimates. With the exception of the latter effect, the model gives a good quantitative account of sources of wage volatility in this decomposition. This can be seen visually in Figure 5.

As we mentioned, changes in the transition rates have direct effects on the wage distribution and indirect effects through the reservation wage. We perform two alternative simulation experiments to demonstrate that it is movements in the reservation wage that are driving our results. Let $\{\lambda_t, \delta_t, x_t^*\}_{t=0}^T$ be the offer arrival rates, separation rates and reservation wages from our simulation from 1948 to 1992 and let λ , δ , and x^* be their steady state values. First, we simulate the model

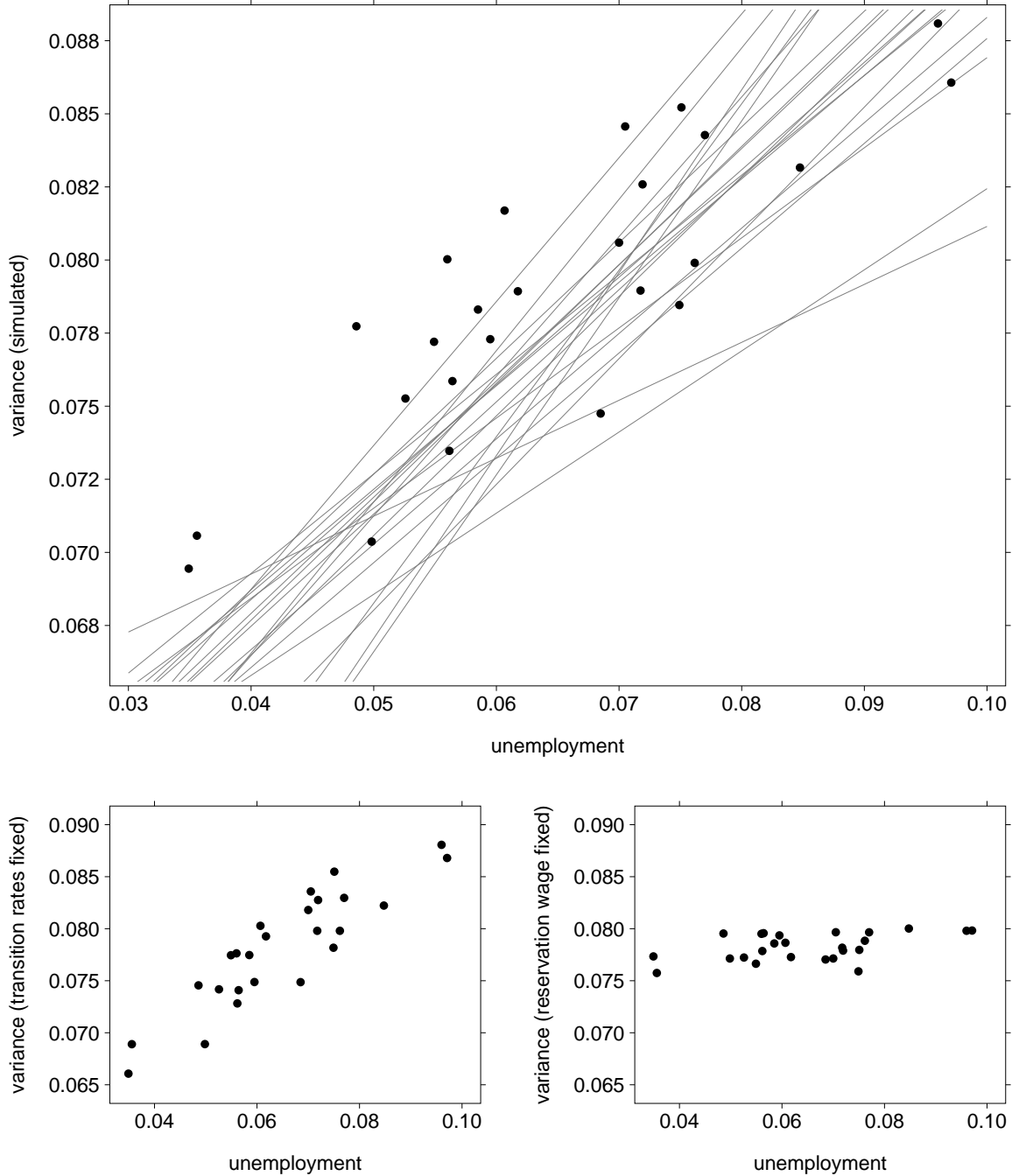


Figure 6: Variance of log wage changes vs unemployment rate for simulated model. Top: variances from simulated model, displayed with posterior draws from posterior of the relevant multilevel regression lines of the empirical estimates in Section 2. Note how the model is able to replicate a significant part of the cyclical variation in the data, but the simulated variance for low values of unemployment is somewhat higher than in the data. Bottom left: simulated series with transition rates fixed at their means, time-varying reservation wage same as the calibrated simulation. Bottom right: reservation wages fixed, transition rates changing. Notice how the the reservation wages generate almost all the variation, not the transition rates.

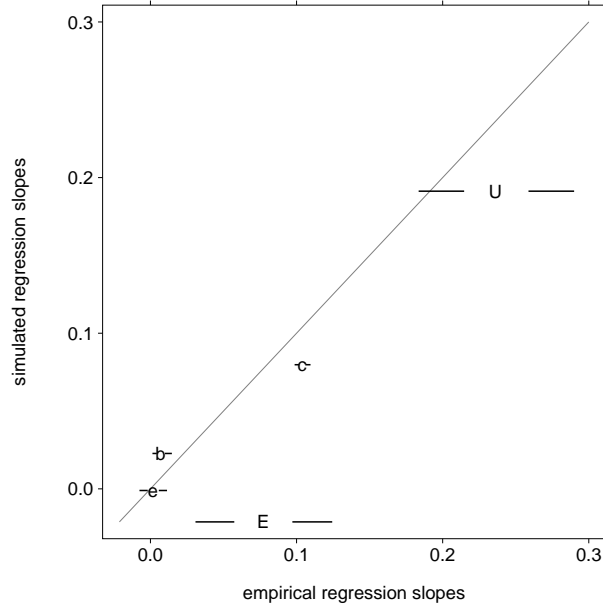


Figure 7: Estimated and simulated decomposition — visualization of Table 5. The horizontal axis shows the empirical quantiles (using the same convention as the rest of the graphs, 5-25-50-75-95% quantiles), while the vertical axis shows the simulated decomposition. The letters in the middle mark the median, with the following legend: (U)nemployed group, no un(E)mployment group, (c)omposition, (b)etween, (e)rror. The grey line is the 45° diagonal. Note that the model comes close to matching all parts of the decomposition, with the exception of group E.

with the transition rates held constant at their steady state values, but assuming that workers use the reservation wage x_t^* in period t . The results appear in the bottom left panel of Figure 6. The slope of a linear regression fit to these data is 0.319. Second, we conduct the opposite simulation exercise: holding the reservation wage fixed at its steady state value and allowing the transition rates to fluctuate. The results appear in the bottom right panel of Figure 6 and the slope of a linear regression in these data is 0.036. From these simulations, it is clear that movements in the reservation wage are the main source of counter-cyclical wage volatility in our model. An increase in the reservation wage reduces the dispersion of wage growth rates because when workers transit through unemployment they re-enter employment at a narrower range of wages. Over the cycle, increases in the offer arrival rate and decreases in the separation rate increase the value of searching and raise the reservation wage. Therefore the reservation wage is pro-cyclical, which induces counter-cyclical movements in the dispersion of wage growth rates.

These alternative simulations also illustrate why the model cannot match the counter-cyclical

wage volatility among those not experiencing unemployment. Those who do not go through unemployment are unaffected by changes in the reservation wage and as a result their wage dynamics are only affected by the direct effect of changes in λ_t and δ_t , which is a small effect.

The model generates a plausible level of volatility in the reservation wage. Figure 8 shows the reservation wage relative to its the steady state value. For most of our sample period, the reservation wage is within 15% of the steady state value. As a final check on our analysis, Figure 9 verifies that steady state reservation wage, which we analyzed in section 4.2, gives a good approximation to the behavior of the reservation wage in the full dynamic model.

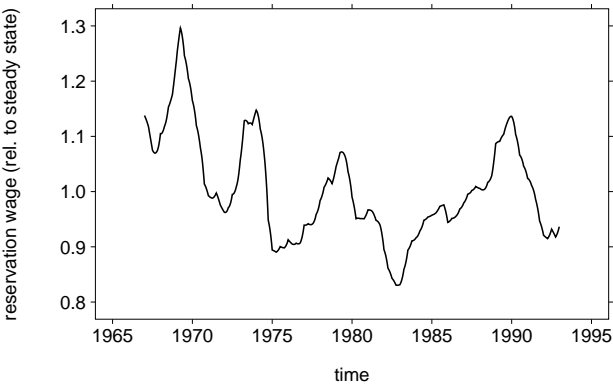


Figure 8: Reservation wage relative to steady state value.

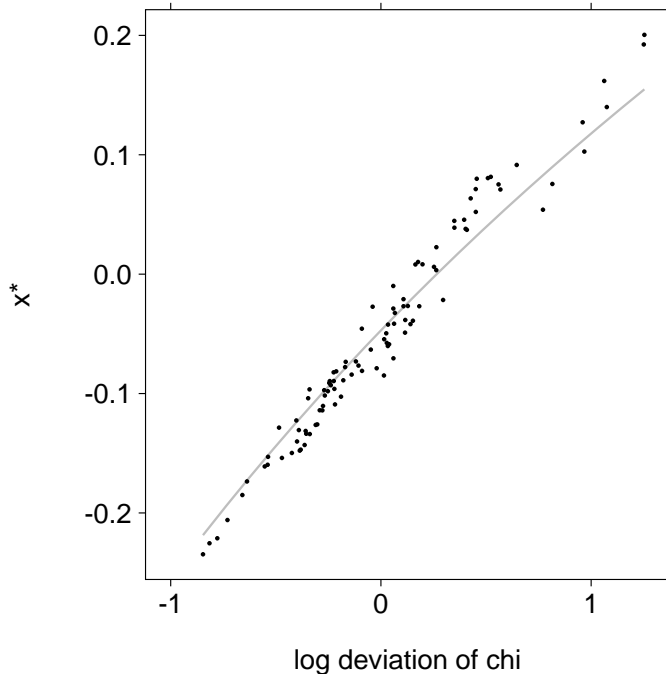


Figure 9: The dots show the (log) reservation wage from the full dynamic model in Section 4.4. The line shows the steady state reservation wage implied by inserting contemporaneous values of λ and δ into equation (20). This graph demonstrates that (1) χ is indeed a good summary of the labor market conditions, and (2) steady state comparative statics provides a good approximation to the behavior of the full dynamic model.

5 Conclusion

In this paper, we combine empirical and theoretical approaches to examine cyclical movements in wage risk. We document that wage volatility increases with the unemployment rate and then decompose this finding to investigate its sources. We find that the composition effect resulting from the greater incidence of unemployment during a recession explains about 25% of the cyclical variation in wage volatility and a further 55% is explained by the cyclical changes in the volatility of wages among those experiencing unemployment.

A standard, search-theoretic model of the labor market can quantitatively explain these effects. In the model, the most important source of cyclical changes in wage volatility among agents going through unemployment is the response of the reservation wage to labor market conditions. We explore this response both analytically via steady state comparative statics and through dynamic

simulation. We find that our calibrated model gives a good account of the cyclical variation in idiosyncratic wage risk among those experiencing unemployment and of the composition effect over the business cycle. As our empirical estimates imply these components generate around 80% of the total cyclical variation in wage volatility, we argue that such a model gives a good account of cyclical variation in idiosyncratic wage risk.

Our findings suggest that search-and-matching models are a promising way forward in incorporating structural models of the labor market into models of household heterogeneity. However, we have focused our analysis on an important, but specific, feature of the wage process. As we noted in the introduction, the persistence of labor market shocks is an important determinant of a household's ability to self insure against idiosyncratic shocks and further research is needed to determine the persistence of the additional shocks that occur in recessions and to assess the performance of search-and-matching models on this dimension.

Finally, multilevel Bayesian methods provide a natural way of documenting business cycle facts from microeconomic data and complement variance decomposition techniques. The methods we have used here can be applied to different cross-sectional moments and different partitions of the data to investigate other distributional effects of business cycle fluctuations.

References

- Atkeson, A. and Phelan, C. (1994). Reconsidering the costs of business cycles with incomplete markets. In *NBER Macroeconomics Annual 1994, Volume 9*, NBER Chapters, pages 187–218. National Bureau of Economic Research, Inc.
- Beaudry, P. and Pages, C. (2001). The cost of business cycles and the stabilization value of unemployment insurance. *European Economic Review*, 45(8):1545–1572.
- Bils, M., Chang, Y., and Kim, S.-B. (2011). Worker heterogeneity and endogenous separations in a matching model of unemployment fluctuations. *American Economic Journal: Macroeconomics*, 3(1):128–54.
- Blau, D. M. and Robins, P. K. (1990). Job search outcomes for the employed and unemployed. *Journal of Political Economy*, 98(3):pp. 637–655.
- Brooks, S. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455.
- Constantinides, G. M. and Duffie, D. (1996). Asset pricing with heterogeneous consumers. *Journal of Political Economy*, 104(2):219–40.
- Costain, J. and Reiter, M. (2008). Business cycles, unemployment insurance, and the calibration of matching models. *Journal of Economic Dynamics and Control*, 32(4):1120–1155.
- Costain, J. S. and Reiter, M. (2005). Stabilization versus insurance: Welfare effects of procyclical taxation under incomplete markets. Technical report, Department of Economics and Business, Universitat Pompeu Fabra.
- Costain, J. S. and Reiter, M. (2007). On the relevance of saving for labor market fluctuations. Technical report, Bank of Spain and Institute for Advanced Studies.
- De Santis, M. (2007). Individual consumption risk and the welfare cost of business cycles. *American Economic Review*, 97(4):1488 – 1506.

- Fallick, B. and Fleischman, C. A. (2004). Employer-to-employer flows in the u.s. labor market: the complete picture of gross worker flows. Finance and Economics Discussion Series 2004-34, Board of Governors of the Federal Reserve System (U.S.).
- Gelman, A. (2005). Analysis of variance — why it is more important than ever. *The Annals of Statistics*, 33(1):1–53.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian analysis*, 1(3):515–533.
- Gelman, A., Carlin, B., Stern, H., and Rubin, D. (2004). *Bayesian Data Analysis*. Chapman and Hall, 2nd edition.
- Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Gelman, A. and Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472.
- Gomes, J., Greenwood, J., and Rebelo, S. (2001). Equilibrium unemployment. *Journal of Monetary Economics*, 48(1):109–152.
- Heathcote, J., Perri, F., and Violante, G. L. (2010). Unequal we stand: An empirical analysis of economic inequality in the united states: 1967-2006. *Review of Economic Dynamics*, 13(1):15–51.
- Hornstein, A., Krusell, P., and Violante, G. (2012). Frictional wage dispersion in search models: a quantitative assessment. *American Economic Review*, page forthcoming.
- Imrohoroglu, A. (1989). Cost of business cycles with indivisibilities and liquidity constraints. *Journal of Political Economy*, 97(6):1364–83.
- Juhn, C., Murphy, K. M., and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3):410–42.
- Jung, P. and Kuester, K. (2011). The (un)importance of unemployment fluctuations for the welfare cost of business cycles. *Journal of Economic Dynamics and Control*, 35(10):1744 – 1768.

- Krebs, T. (2003). Growth and welfare effects of business cycles in economies with idiosyncratic human capital risk. *Review of Economic Dynamics*, 6(4):846–868.
- Krebs, T. (2007). Job displacement risk and the cost of business cycles. *American Economic Review*, 97(3):664–686.
- Krueger, A. B. and Mueller, A. (2011). Job search and job finding in a period of mass unemployment: Evidence from high-frequency longitudinal data. Technical Report 215, Princeton University Center for Economic Policy Studies.
- Krueger, D. and Lustig, H. (2010). When is market incompleteness irrelevant for the price of aggregate risk (and when is it not)? *Journal of Economic Theory*, 145(1):1 – 41.
- Krusell, P., Mukoyama, T., Rogerson, R., and Sahin, A. (2011). Is labor supply important for business cycles? Technical report, University of Virginia.
- Krusell, P., Mukoyama, T., and Sahin, A. (2010). Labour-market matching with precautionary savings and aggregate fluctuations. *Review of Economic Studies*, 77(4):1477–1507.
- Krusell, P., Mukoyama, T., Sahin, A., and Smith, A. A. (2009). Revisiting the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 12(3):393–402.
- Krusell, P. and Smith, A. A. (1999). On the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 2(1):245–272.
- Lange, K. (2010). *Applied probability*. Springer Verlag.
- Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *The American Economic Review*, 96(3):461–498.
- Lucas, R. (1987). *Model of Business Cycles*. Blackwell.
- Lucas, R. J. and Prescott, E. C. (1974). Equilibrium search and unemployment. *Journal of Economic Theory*, 7(2):188–209.

- Mankiw, N. G. (1986). The equity premium and the concentration of aggregate shocks. *Journal of Financial Economics*, 17:211–19.
- McCall, J. J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–26.
- Menzio, G. and Shi, S. (2011). Efficient search on the job and the business cycle. *Journal of Political Economy*, 119(3):468 – 510.
- Nagypál, É. (2005). On the extent of job-to-job transitions.
- Nakajima, M. (2010). Business cycles in the equilibrium model of labor market search and self-insurance. Working Papers 10-24, Federal Reserve Bank of Philadelphia.
- Neal, R. (2003). Slice sampling. *Annals of Statistics*, 31(3):705–741.
- Polson, N. and Scott, J. (2011). On the half-cauchy prior for a global scale parameter. *Arxiv preprint arXiv:1104.4937*.
- Rudanko, L. (2009). Labor market dynamics under long-term wage contracting. *Journal of Monetary Economics*, 56(2):170–183.
- Rudanko, L. (2011). Aggregate and idiosyncratic risk in a frictional labor market. *American Economic Review*, 101(6):2823–43.
- Shao, E. and Silos, P. (2007). Uninsurable individual risk and the cyclical behavior of unemployment and vacancies. Working Paper 2007-05, Federal Reserve Bank of Atlanta.
- Shimer, R. (2005). The cyclical behavior of unemployment and vacancies. *American Economic Review*, 95(1):25–49.
- Shimer, R. (2007). Reassessing the ins and outs of unemployment. Technical report, University of Chicago.
- Sims, C. (2007). Bayesian methods in applied econometrics, or, why econometrics should always and everywhere be bayesian. *Hottelling lecture, June*, 29:2007.

- Storesletten, K., Telmer, C., and Yaron, A. (2007). Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics*, 10(4):519–548.
- Storesletten, K., Telmer, C. I., and Yaron, A. (2001). The welfare cost of business cycles revisited: Finite lives and cyclical variation in idiosyncratic risk. *European Economic Review*, 45(7):1311–1339.
- Storesletten, K., Telmer, C. I., and Yaron, A. (2004). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*, 112(3):695–717.
- Western, B. and Bloome, D. (2009). Variance function regressions for studying inequality. *Sociological Methodology*, 39(1):293–326.

Appendix

A Data Appendix

From the PSID, we use data on the annual labor income of male heads and annual hours of work to construct annual wages. In addition, we use data on age and education as covariates. We only include those heads that are between 25 and 60 years of age in both years over which the change in wages is calculated. We drop those with allocated labor income, students, business owners, self-employed individuals, and those with zero hours or income. We trim the top 1% of the income distribution in each year to remove the effect of changes in top-codes across years. Finally, we drop those with wages less than half of the federal minimum wage in that year and we drop those who work fewer than 320 hours in a given year. Our results on income changes in section 2.4 are based on the same sample, but we use annual labor income without dividing by annual hours.

To construct the unemployment-experience partition, we use data on annual hours of unemployment. Let H_t^U be the hours of unemployment in year t . The unemployment group at time t includes those individuals who report $H_t^U > 0$ or $H_{t-1}^U > 0$. The no-unemployment group is those individuals for whom H^U is equal to zero for both years. For aggregate data, we use the national unemployment rate reported by the BLS and take the average value of the monthly series within each year. Figure 10 shows the share of the unemployment group in our sample.

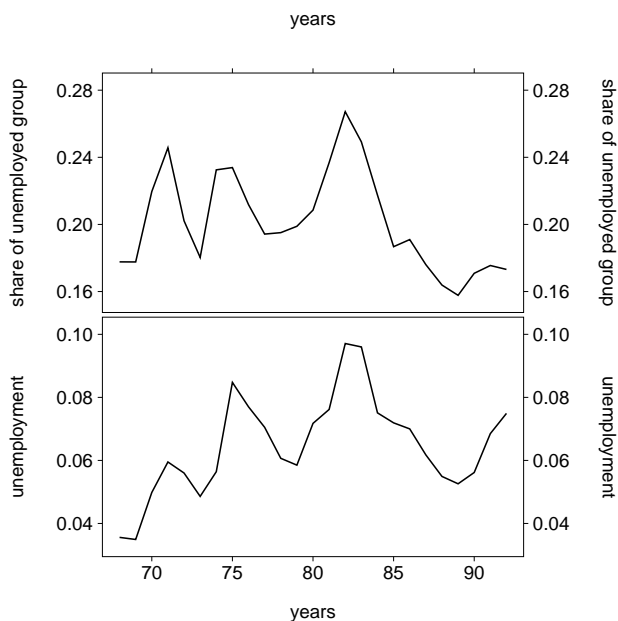


Figure 10: Share of observations in the “unemployment” group (top panel) displayed along with the unemployment rate (bottom panel). As expected, unemployment incidence is highly correlated with the unemployment rate.

Note: the rest of the appendix is available online, from the authors’ webpages.

Note: the remaining appendices are for online distribution only.

B Methodological Appendix

B.1 The Gibbs sampler

Sampling from the posterior of our model is straightforward and follows Markov Chain Monte Carlo techniques that are commonly used in Bayesian statistics. There are several alternative methods for sampling from the posterior (see Gelman et al. (2004) for a summary). We found that a block Gibbs sampler is relatively fast, easy to implement and has good mixing properties (see Section B.2 for a discussion of convergence). A brief summary of the Gibbs sampler we used is given below.

Our model can be summarized by equations (14)–(16). This model collapses to that in section 2 if one sets $J = 1$ with $j[i, t] = 1$ for all i and t . We partition the parameter space into blocks corresponding to θ , α , κ , ς_α , σ^2 , η , and ς_{σ^2} .

B.1.1 Conditional posterior sampling for θ , α , κ , and η

Each of these parameters can be thought of as the coefficient v in a linear regression

$$c \sim N(Av, \Phi) \tag{24}$$

with a known variance matrix $\Phi = \text{diag}(\phi)$ and a given prior $v \sim N(\Pi)$. Table 6 shows the mapping between (24) and the model parameters. The conditional posterior distribution is normal, and can be constructed and sampled from in a straightforward manner (Gelman et al., 2004, Sections 14.6 and 14.8).

v	c	A	ϕ	Π
θ	$[dw_{i,t} - \alpha_{t,j[i,t]}]_{i,t}$	$[X_{i,t}]_{i,t}$	$[\sigma_{t,j[i,t]}^2]_{i,t}$	flat
α_j	$[dw_{i,t} - X_{i,t}\theta]_{i,t:j[i,t]=j}$	$\mathbf{1}$	$[\sigma_{t,j[i,t]}^2]_{i,t:j[i,t]=j}$	ς_α
κ_j	$\alpha_{\cdot,j}$	Z	$\varsigma_{\alpha j}$	flat
η_j	$\sigma_{\cdot,j}^2$	Z	$\varsigma_{\sigma^2 j}$	flat

Table 6: Conditional posterior sampling for θ , α , κ , and η

B.1.2 Conditional posterior sampling for the group-level variances ς_α and ς_{σ^2}

As is well known, if $h_k \sim N(0, \varsigma^2)$ (iid, $k = 1, \dots, K$), then the likelihood is

$$p(\varsigma | h) \propto \varsigma^{-K} \exp\left(-\frac{\sum_{k=1}^K h_k^2}{2\varsigma^2}\right) \tag{25}$$

As discussed above (see Section 2.1.1), our prior for hyperparameter variances is $p(\varsigma) \propto (\varsigma^2 + s^2)^{-1}$. Then we sample from the conditional posterior using the slice sampler of Neal (2003). Table 7 shows the correspondences between (25) and the model parameters.

h	ς	s
$\alpha_{\cdot,j} - Z\kappa_j$	$\varsigma_{\alpha j}$	s_{α}
$\sigma_{\cdot,j}^2 - Z\eta_j$	$\varsigma_{\sigma^2 j}$	s_{σ^2}

Table 7: Conditional posterior sampling for ς_{α} , ς_{σ^2}

B.1.3 Conditional posterior sampling for σ^2

Let us fix j and t , and only consider the observations in $I_{t,j} = \{(i', t') : t' = t, j[i', t'] = j\}$. Given (16) and (8), the conditional posterior for $\sigma_{j,t}^2$ is

$$p(\sigma_{t,j}^2 \mid \alpha_{t,j}, Z_t, \eta_j, \varsigma_{\sigma^2}) \propto (\sigma_{t,j}^2)^{-\frac{|I_{j,t}|}{2}} \exp\left(-\frac{\sum_{(i,t) \in I_{t,j}} (dw_{i,t} - X_{i,t}\theta - \alpha_{t,j})^2}{2\sigma_{t,j}^2}\right) \exp\left(-\frac{(\sigma_{t,j}^2 - Z\eta_j)^2}{2\varsigma_{\sigma^2 j}}\right) 1_{\sigma_{t,j}^2 \geq 0} \quad (26)$$

This does not correspond to any commonly used family of probability distributions, so we can only sample from it using general tools. After experimenting with rejection methods and obtaining poor acceptance rates, we settled on the slice sampling algorithm of Neal (2003) with excellent results.¹⁸

B.2 Convergence of the posterior sampler

We monitor the convergence of the Gibbs sampler by calculating the univariate potential scale reduction factor (PSRF) for each parameter value (Gelman and Rubin, 1992; Brooks and Gelman, 1998; Gelman et al., 2004). The PSRF uses variances within and between the parallel chains to estimate the factor by which the scale of the current posterior distribution for a given parameter might be reduced if we were to obtain a sample of infinite size. In practice, values below 1.1 are acceptable, unless very high precision is required. We start 5 chains with overdispersed initial points, calculate the PSRF as the chain evolves for the second half of the chain, and find that it goes below 1.01–1.05 (depending on the parameter) for all parameters after 1000–2000 iterations, which suggests that the mixing is excellent. We generate 5000 parameters for each chain, and discard the first 2500. We found that the mixing was greatly enhanced by subtracting the column means from X .

B.3 The non-hierarchical model

We estimate a version of the model without hierarchical regressions for comparison.

$$\begin{aligned} dw_{i,t} &\sim N(X_{i,t}\theta + \alpha_{t,j[i,t]}, \sigma_{t,j[i,t]}^2) \\ p(\theta) &\propto 1 \\ p(\alpha_{t,j}) &\propto 1 \quad \forall t, j \\ p(\sigma_{t,j}^2) &\propto (\sigma_{t,j}^2)^{-1} \quad \forall t, j \end{aligned}$$

¹⁸In particular, we used the “stepping out” variant of the algorithm from Neal (2003, Figures 3 and 5).

It is very straightforward to sample from the posterior of this model: first we sample $\theta \mid \alpha, \sigma^2, X$ (see Section B.1.1), then sample from $\alpha_{t,j}, \sigma_{t,j}^2 \mid \theta, X$ by sampling from the posterior of the regression of $dw_{i,t} - \alpha_{t,j[i,t]}$ on 1 with unknown variance and a reference prior for each t, j .

C Comparison to Storesletten et al. (2004)

The purpose of this appendix is to express Storesletten et al.'s results in the same terms as ours to show that the two are not vastly at odds with one another. Throughout, we take their results reported in their Table 2, Panel E because these results are calculated on the assumption that business cycles are defined by the unemployment rate as opposed to GNP growth or NBER cycles and thus closest to our work.

Storesletten et al. specify the following process for the residual of log earnings of individual i

$$\begin{aligned} u_{it} &= \alpha_i + z_{it} + \varepsilon_{it} \\ z_{it} &= \rho z_{i,t-1} + \eta_{it}, \end{aligned}$$

where $\alpha_i \sim \text{Niid}(0, \sigma_\alpha^2)$, $\varepsilon_{it} \sim \text{Niid}(0, \sigma_\varepsilon^2)$, and $\eta_{it} \sim \text{Niid}(0, \sigma_E^2)$ in an expansion and $\eta_{it} \sim \text{Niid}(0, \sigma_C^2)$ in a contraction. In this notation, our interest is in computing the variance of Δu_{it} , which is

$$\text{Var}[\Delta u_{it}] = (\rho - 1)^2 \text{Var}[z_{i,t-1}] + \text{Var}[\eta_{it}] + \text{Var}[\Delta \varepsilon_{it}]$$

Clearly, as $\rho \rightarrow 1$, the first term on the right-hand side goes to zero. This is relevant because Storesletten et al. estimate ρ to be close to 1. For now, suppose $\rho = 1$, but we return to the issue below. Since ε is distributed identically over time, the $\text{Var}[\Delta \varepsilon_{it}]$ term is a constant. Thus, the difference in $\text{Var}[\Delta u_{it}]$ between an expansion and a contraction is just the difference in $\text{Var}[\eta_{it}]$ or $\sigma_C^2 - \sigma_E^2$. Using Storesletten et al.'s estimates, this difference is $0.246^2 - 0.138^2 = 0.041$. In our back of the envelope calculation in section 2.4 we found a difference of 0.013 for wages and 0.034 for earnings.

In the calculation above we ignored the term $(\rho - 1)^2 \text{Var}[z_{i,t-1}]$, which will be counter-cyclical as the variance of z is counter-cyclical. As argued above, this term is small. To see this, consider two extreme economies, one that is always in an expansion and one that is always in a contraction. The unconditional variance of z in the expanding economy is then $\sigma_E^2 / (1 - \rho^2)$ and one can similarly calculate the unconditional variance of z in the contracting economy. The difference between these two variances is an upper bound on the cyclical fluctuations in $\text{Var}(z)$. Using this upper bound, we conclude that the contribution of the term $(\rho - 1)^2 \text{Var}[z_{i,t-1}]$ is at most 0.0013 or 3% of the 0.041 figure we found above.

D Proofs for the steady state analysis

D.1 Linear separability and the reservation wage

Using the independence of the stochastic processes, we rewrite (17) and (18) in terms of

$$\begin{aligned} W(p, a, x) &= W(x) + H(p, a) \\ U(p, a) &= U + H(p, a) \end{aligned}$$

where

$$H(p, a) = p + a + \beta \mathbb{E}_{(p', a')} [H(p', a')]$$

and

$$U = b + \lambda \beta \mathbb{E}_{x'} \max[W(x'), U] \quad (27)$$

$$W(x) = x + \beta (\gamma \lambda \mathbb{E}_{x'} \max[W(x'), W(x)] + \delta U + (1 - \delta - \gamma \lambda) W(x)) \quad (28)$$

Clearly, W is increasing and the unemployed worker follows a reservation rule. Let x^* be the reservation productivity, defined by

$$W(x^*) = U \quad (29)$$

Integration by parts yields¹⁹

$$\mathbb{E}_{x'} \max[W(x'), W(x)] - W(x) = \int_x^\infty W'(x') \bar{F}(x') dx'$$

Then partial differentiation of (28) and the combination of (27) and (28) evaluated at $x = x^*$ gives us (19).

D.2 Proof of Lemma 1

Instead of (20), we work with the equivalent equation

$$x^* = b + \int_{x^*}^\infty \frac{(1 - \gamma) \chi \bar{F}(x)}{1 + \gamma \chi \bar{F}(x)} dx \quad (30)$$

Total differentiation of (30) and a bit of algebra gives us

$$dx^* = \hat{d}\chi \frac{\chi}{1 + \chi \bar{F}(x^*)} \underbrace{(1 + \gamma \chi \bar{F}(x^*)) (1 - \gamma) \int_{x^*}^\infty \frac{\bar{F}(x)}{(1 + \chi \gamma \bar{F}(x))^2} dx}_{\mathcal{Q}(\gamma)} \quad (31)$$

¹⁹We assume that $\lim_{x \rightarrow \infty} \bar{W}(x) F(x) = 0$. Since $W(x)$ is concave, the existence of the first moment of F is sufficient for this to happen.

We derive a first-order approximation for the part denoted $\mathcal{Q}(\gamma)$ by perturbing around $\gamma = 0$:

$$\begin{aligned}\mathcal{Q}(0) &= \int_{x^*}^{\infty} \bar{F}(x) dx \\ \mathcal{Q}'(0) &= (\chi \bar{F}(x^*) - 1) \int_{x^*}^{\infty} \bar{F}(x) dx - 2\chi \int_{x^*}^{\infty} \bar{F}(x)^2 dx \\ \mathcal{Q}(\gamma) &\approx \left(\int_{x^*}^{\infty} \bar{F}(x) dx \right) \left(1 + \gamma \left(\chi \bar{F}(x^*) - 1 - 2\chi \frac{\int_{x^*}^{\infty} \bar{F}(x)^2 dx}{\int_{x^*}^{\infty} \bar{F}(x) dx} \right) \right)\end{aligned}$$

Finally, (21) obtains from substituting the expansion above into (31), collecting terms and rearranging.

D.3 Closed-form expressions for the tail characteristics of the standard normal distribution

Here, we derive a closed-form solution for M_γ when the offer distribution is normal. As discussed in footnote 14, it is sufficient to focus on the standard normal. We use Φ for the CDF of the standard normal. Let $\bar{\Phi}(z) = 1 - \Phi(z)$, and $\phi(z) = \Phi'(z) = -\bar{\Phi}'(z)$. First, we establish the following limit: for all $a \in \mathcal{R}$,

$$\bar{\Phi}(a) = \frac{1}{\sqrt{2\pi}} \int_a^{\infty} e^{-\frac{z^2}{2}} dz < \frac{1}{\sqrt{2\pi}} \int_a^{\infty} \frac{z}{a} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \frac{1}{a} e^{-\frac{a^2}{2}}$$

and thus

$$\lim_{z \rightarrow \infty} \bar{\Phi}(z)z = 0 \tag{32}$$

Then using the transformation²⁰

$$\begin{aligned}\int_{z^*}^{\infty} \bar{\Phi}(z) dz &= \underbrace{[\bar{\Phi}(z)(z - z^*)]_{z^*}^{\infty}}_{=0, \text{ because of (32)}} + \int_{z^*}^{\infty} \phi(z)(z - z^*) dz = \int_{z^*}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} z dz - z^* \bar{\Phi}(z^*) \\ &= \phi(z^*) - z^* \bar{\Phi}(z^*)\end{aligned}$$

we obtain the following closed-form solution:

$$M_1(z^*) = \frac{\phi(z^*)}{\bar{\Phi}(z^*)} - z^* \tag{33}$$

and similarly,

$$\begin{aligned}\int_{z^*}^{\infty} \bar{\Phi}^2(z) dz &= \underbrace{[\bar{\Phi}(z)(z - z^*)]_{z^*}^{\infty}}_{=0, \text{ because of (32)}} + \int_{z^*}^{\infty} 2\bar{\Phi}(z)\phi(z)(z - z^*) dz \\ &= -z^* \bar{\Phi}^2(z^*) + 2\bar{\Phi}(z^*)\phi(z^*) - \int_{z^*}^{\infty} 2\phi^2(z) dz = -z^* \bar{\Phi}^2(z^*) + 2\bar{\Phi}(z^*)\phi(z^*) - \pi^{-1/2} \bar{\Phi}(\sqrt{2}z^*)\end{aligned}$$

²⁰The result also follows from Stein's Lemma, eg Lange (2010, p 40).

can be used to show that

$$M_2(z^*) = 2 \frac{\phi(z^*)}{\bar{\Phi}(z^*)} - \pi^{-1/2} \frac{\bar{\Phi}(\sqrt{2}z^*)}{\bar{\Phi}^2(z^*)} - z^*. \quad (34)$$

Finally, M_γ is given by $2M_2(z^*)/M_1(z^*) - 1$.

E Additional discussion of calibration

E.1 Calibration of the stochastic process for a_t

According to our model, in year t , the log wage of a continuously employed individual is given by

$$w_t = \log \left[\frac{1}{48} \sum_{s=1}^{48} \exp(a_t(s) + x_t(s)) \right] + \epsilon_t,$$

where s indexes the weeks within year t and ϵ_t is measurement error. If we restrict our attention to those on the same job, then the value of $x_t(s)$ is constant within and across years. As a result, the change in the log wage is given by

$$\Delta w_t = \Delta \log \left[\frac{1}{48} \sum_{s=1}^{48} \exp(a_t(s)) \right] + \Delta \epsilon_t.$$

We will write this as

$$\Delta w_t = \Delta A_t + \Delta \epsilon_t,$$

where A_t is the log-sum term above. A_t will follow a stationary stochastic process that depends on the underlying AR(1) for the a process with parameters ρ and σ .

The autocovariances of Δw_t are

$$\begin{aligned} \text{Var}(\Delta w_t) &= 2 \text{Var}(A_t) - 2 \text{Cov}(A_t, A_{t-1}) + 2 \text{Var}(\epsilon_t) \\ \text{Cov}(\Delta w_t, \Delta w_{t-1}) &= -\text{Var}(A_t) - \text{Cov}(A_t, A_{t-2}) - \text{Var}(\epsilon_t) \\ \text{Cov}(\Delta w_t, \Delta w_{t-2}) &= 2 \text{Cov}(A_t, A_{t-2}) - \text{Cov}(A_t, A_{t-1}) - \text{Cov}(A_t, A_{t-3}). \end{aligned}$$

We simulate the process $a_t(s)$ and use it to construct A_t . To calibrate the model, we choose ρ , σ , and $\text{Var}(\epsilon_t)$ to minimize the sum of squared differences between the simulated autocovariances and the same moments computed from the data. To compute the moments in the data, we use our same PSID sample,²¹ but restrict our attention to individuals with levels of tenure on the job high enough that the differences in wages and the covariances across years involve wages from the same job. For instance, $\text{Cov}(\Delta w_t, \Delta w_{t-2})$ involves w_t, \dots, w_{t-3} and so we require that the only individuals with job tenure of at least $48 + [\text{interview month}]$ months enters into this calculation. The results are $\text{Var}(\Delta w_t) = .04623$, $\text{Cov}(\Delta w_t, \Delta w_{t-1}) = -.016702$, and $\text{Cov}(\Delta w_t, \Delta w_{t-2}) = -.000987$. We also find the unconditional variance of wage growth rates to be 0.078037.

²¹In a first-stage regression we remove age, education and year effects as we do in Sections 2 and 3.

E.2 Calibration of the relative search efficiency γ

Nagypál (2005) shows that in wage-ladder models it is possible to calibrate the relative search efficiency (or equivalently, the job finding rate) of employed workers merely from the observed transition rates, *without specifying the offer distribution*. We give a concise summary of the method here and show how it applies in our model.

In steady state, the unemployment rate satisfies

$$\lambda\bar{F}(x^*)u = \delta(1 - u) \quad (35)$$

where the left and right hand sides are the mass of workers matched or separated each period, respectively. Let $G(x)$ denote the cross-sectional fraction of employed workers with match-specific productivity below x . Then flow balance requires that

$$\lambda(F(x) - F(x^*))u = G(x)(\delta + \gamma\lambda\bar{F}(x))(1 - u) \quad (36)$$

The left hand side is the flow of workers from unemployment to employment below a particular x — we need to subtract $F(x^*)$ because no offers are accepted below the reservation x^* . The right hand side accounts for exogenous separations and upward transitions on the wage ladder.

We introduce

$$\varphi \equiv \gamma \frac{\lambda\bar{F}(x^*)}{\delta} = \gamma \frac{\lambda_{UE}}{\delta} \quad (37)$$

which is equal to the ratio of the job finding and separation probabilities, corrected by γ , the relative search efficiency of the employed. All but the last are directly observable from flow data. Below we calibrate φ (and this γ) in a manner independent of the actual distributions F or G , or the reservation x^* . Then we rewrite (36) using (35), (37) and the normalized tail distribution \bar{F}^* (defined in (1)) as

$$1 - \bar{F}^*(x) = G(x)(1 + \varphi\bar{F}^*(x))$$

which we can rearrange as

$$G(x) = \frac{1 - \bar{F}^*(x)}{1 + \varphi\bar{F}^*(x)} \quad (38)$$

It follows immediately that the cross-sectional distribution of individual-specific productivities only depends on the transition rates via φ , which we will calibrate directly from flows. It is useful to rearrange (38) as²²

$$\varphi\bar{F}^*(x) = \frac{1 + \varphi}{1 + \varphi G(x)} - 1 \quad (39)$$

We follow Nagypál (2005) in calibrating φ directly. Consider the *average* job-to-job transition probability

$$\lambda_{EE} = \gamma\lambda \int_{x^*}^{\infty} \bar{F}(x) dG(x)$$

²²It is also easy to obtain the inverse from (38): when $G(x) = y$,

$$\bar{F}^*(x) = \frac{1 - y}{1 + \varphi y} \quad \text{and thus} \quad x = \bar{F}^{*-1} \left(\frac{1 - y}{1 + \varphi y} \right)$$

which is useful for initializing the simulation.

which we can rewrite as

$$\frac{\lambda_{EE}}{\delta} = \int_{x^*}^{\infty} \varphi \bar{F}^*(x) dG(x) = \int_{x^*}^{\infty} \frac{1 + \varphi}{1 + \varphi G(x)} dG(x) - 1$$

Using the substitution $y = G(x)$, we arrive at

$$\frac{\lambda_{EE}}{\delta} = \frac{1 + \varphi}{\varphi} \log(1 + \varphi) - 1 \tag{40}$$

where the left hand side is observable from flow data, and the right hand side is a monotone increasing function of φ , with limits of 0 and ∞ at $\varphi \rightarrow 0$ and $\varphi \rightarrow \infty$, respectively.

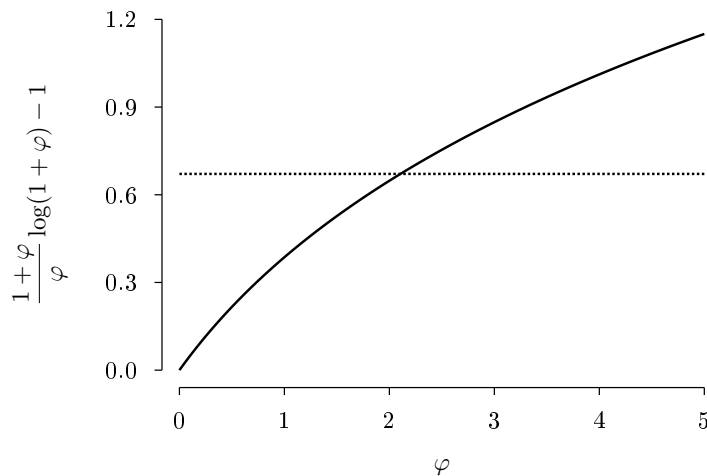


Figure 11: Calibration of φ . The solid line shows the right hand side of (40), while the dotted line shows the target value of λ_{EE}/δ , where we target a monthly job-to-job transition probability of 2.6% (Fallick and Fleischman, 2004), converted to a weekly rate, and the separation rate from Table 4. We solve numerically for $\varphi = 2.106$.