Income Dispersion and Counter-Cyclical Markups

Chris Edmond and Laura Veldkamp*
New York University
Stern School of Business

April 11, 2007

Abstract

Income heterogeneity rises in recessions (Storesletten, Telmer, and Yaron 2004). The counter-cyclical nature of income dispersion makes the optimal markups charged by monopolistically competitive firms rise in recessions. The reason is that when incomes, and thus willingness to pay, are more dispersed, the aggregate elasticity of demand falls. We illustrate this effect in an equilibrium model and quantify its importance for macroeconomic variables such as output, labor supply, and prices.

Keywords: business cycles, counter-cyclical markups, income dispersion.

JEL classifications: E32.

Traditional business cycle models have had some success in explaining the fluctuations in labor, investment and output, using a representative agent framework. But this literature has largely ignored the cyclical behavior of prices. A parallel monetary economics literature focuses on different business cycle facts – smooth prices, procyclical firm profits and counter-cyclical markups – using imperfect competition in conjunction with firm or demand heterogeneity. But this literature does not attempt to match the standard business cycle quantities.[1] Recent research that quantifies income heterogeneity offers a way to connect the ideas of these two traditions. Using estimates of the counter-cyclical behavior of income dispersion, we develop a model that can simultaneously explain key features of price behavior, and business cycle quantities.

Our mechanism is a simple one: Changes in earnings dispersion change the price elasticity of demand over the business cycle. When earnings are more dispersed, buyers' willingness to pay is also more dispersed. If sellers were to reduce prices in recessions, they would attract few additional customers (as in the left panel of figure 1). This low elasticity makes the marginal benefit of lowering prices smaller and induces firms to keep prices high. Therefore, when dispersion is high, prices stay high, but profits are low. In contrast, in booms when dispersion is low, a seller who lowers her price attracts many additional customers (as in the right panel of figure 1). Therefore in booms, sellers keep prices low but earn high profits.

The goal of this paper is two-fold. First, we show how and under what conditions increased earnings dispersion can cause markups to rise. Second, we embed the earnings process estimated by Storesletten, Telmer, and Yaron (2004) in a calibrated business cycle model and show that it can explain the extent of cyclical movements in markups, profits and traditional business cycle quantities. A key challenge in achieving these two goals is to keep

---

[1]Of course, this is a coarse characterization of these two literatures. As always, there are exceptions. Examples include Jaimovich (2006), Lorenzoni (2006), Comin and Gertler (2006), Altig, Christiano, Eichenbaum, and Linde (2004) and a large consumption-based asset pricing literature that uses business cycle models to explain prices of risky claims on capital.
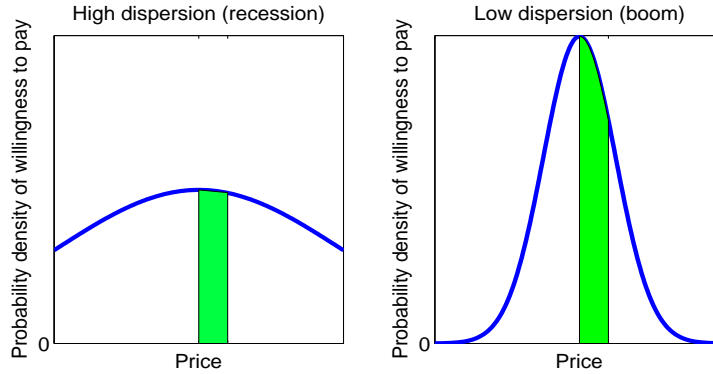
Figure 1: LOWERING PRICE IS MORE BENEFICIAL WHEN DISPERSION IS LOW.
The shaded area represents the increase in the probability of trade from lowering the price, by an amount equal to the width of the shaded area. This higher probability, times the expected gains from trade, is the marginal benefit to reducing the price. Willingness to pay is based on agents' earnings.

the model simple and stylized enough so that the mechanism is transparent while keeping it realistic enough to match the data.

To illustrate the workings of the model's key mechanisms, section 1 analyzes a static version of the model. There is a competitive sector and a monopolistically competitive sector that produces differentiated products. Both produce with intermediate goods, whose only input is labor. Agents choose how many hours to work, how much of the competitive good to consume and how many of the differentiated products to buy. Income dispersion arises because some agents are more productive: They produce more units of the intermediate good per hour. By manipulating the idiosyncratic productivity distribution, we we show that more dispersion results in higher markups and higher prices.

The theory results alone do not rule out the possibility that the variation in earnings dispersion is too small or insufficiently counter-cyclical to generate observed markup or profit fluctuations. Therefore section 2 calibrates and simulates a repeated version of the model. Individuals' earnings processes are taken from the estimates of Storesletten, Telmer, and Yaron (2004). Section 2.2 documents our main results: The model's optimal prices are determined primarily by earnings dispersion. Since measured dispersion is counter-cyclical

and smooth, prices and markups in the model are counter-cyclical and smooth, consistent with the findings of empirical studies in both macroeconomics and industrial organization.

To argue that this is a realistic explanation for a business cycle effect, we show that the model produces reasonably behaved macroeconomic aggregates for other variables as well. Section 2.3 compares the model's predictions for real wages, the share of output firms earn as profits, and labor supply to their empirical counterparts. The model does well in capturing the dynamics of wages and profits, but misses on labor supply. This problem arises because of wealth effects on labor. We propose a solution to counteract those effects.

The addition of heterogeneous income increases the welfare costs of business cycles. The fact that prices rise in downturns makes output fall further. Section 2.5 shows that the welfare costs increase at least 3-fold for the average agent and are orders of magnitude higher for the lowest-income agents. Of course, our model abstracts from important issues considered in the literature on income heterogeneity and welfare ((Krusell and Smith 1998), (Rios-Rull 1996) and (Krueger and Perri 2005)), such as consumption-savings behavior. Nevertheless, our welfare costs are similar.

While our goal is to explore the effect of time-varying earnings dispersion on prices and markups, it is also possible that other cyclical mechanisms are generating this effect. One such alternative is that there are simply sticky prices and procyclical marginal costs. Therefore, the difference between price and cost, the markup, is counter-cyclical. The problem with this simple explanation is that it implies counter-cyclical firm profits, strongly at odds with the data. Our model delivers the observed procyclical profits. Booms are times when markups are low but volume is high enough to compensate. Another possibility is that firm entry and exit change the degree of market competition and thus the markup (Jaimovich 2006). The constant zero profits in this model are an improvement, but still at odds with data. In Comin and Gertler (2006), the causality is reversed: They use shocks

to markups as the source of business cycle fluctuations. Two closely related models are also about why the elasticity of demand fluctuates over the business cycle. In Gali (1994), this is due to a change in the composition of demand. In Bilbiie, Ghironi, and Melitz (2006), changing elasticity comes from a change in the number of available products.

To argue that earnings dispersion is at least part of the reason for price variation, we look for other evidence that long-run changes and cross-sectional differences in earnings dispersion is correlated with differences in prices, output volatility, and profit shares, as predicted by the model. Section 3 shows that the observed decline in earnings dispersion can account for a sizeable part of the observed decline in business cycle volatility, the slow-down in real wage growth, and the accompanying increase in profit shares. Section 4.1 uses state-level panel data to test the model's predicted relationships between earnings dispersion and prices. Section 4.2 documents additional facts from the empirical pricing literature that when the customer base has a more dispersed earnings, prices tend to be higher.

Our explanation raises an obvious question: Why does earnings dispersion rise in recession? One explanation is that job destruction in recessions is responsible (Caballero and Hammour 1994). Rampini (2004) argues that entrepreneurs' incentives change in recessions, making firm outcomes and owners' earnings more risky. Cooley, Marimon, and Quadrini (2004) and Lustig and Van Nieuwerburgh (2005) argue that low collateral values inhibit risk-sharing in recessions. Any one of these explanations could be merged with this model to produce a model whose only driving process is aggregate technology shocks.

1 Static model

There is a continuum of agents indexed by $i \in [0, 1]$, with identical preferences over a numeraire good c_i, labor n_i, and a continuum of differentiated products x_j, indexed by

$j \in [0, 1]$

$$U = \ln(c_i) + \theta(1 - n_i) + \nu \int_0^1 x_{i,j} dj. \tag{1}$$

Each of the differentiated products is indivisible. An agent either buys good j or not, $x_{ij} \in \{0, 1\}$. But the total quantity of x goods consumed can be adjusted by buying more or fewer goods.

Labor is used to make intermediate goods y. Each agent's production is their labor input times their productivity: $n_i w_i$. This intermediate good could be thought of as a physical good, or simply as effective labor that agents supply to firms.

Heterogeneous labor productivity is the source of earnings dispersion. The distribution of productivity is summarized by two parameters, z which shifts the mean and σ the dispersion. We call z aggregate productivity.

$$w_i = z + \sigma e_i \tag{2}$$

The distribution of labor productivity shocks is $e_i \sim g(e)$. We place two restrictions on the distribution function g. First, it is log-concave. This simplifies the comparative statics analysis considerably. Second, all shocks are non-negative, $g(e) = 0, \forall e < 0$. This helps to keep wages from becoming negative and causing problems with undefined utility because of negative consumption.

When we move to the dynamic model, it is the standard deviation σ that will shift over time to match the earnings dispersion facts. Aggregate output of the intermediate good is the integral of individual worker's outputs:

$$y = \int_0^1 w_i n_i di \tag{3}$$

Profit-maximizing firms transform intermediate goods into final c goods and differentiated x_j goods. A continuum of perfectly competitive firms transform intermediate goods 1-for-1

into final c goods. There is another continuum of monopolistic competitors, indexed by j. Each owns a technology to transform intermediate goods into x_j goods: $x_j = y^\alpha$, for $0 < \alpha < 1$. The aggregate resource constraint for transforming intermediate into final goods is

$$\int_0^1 c_i di + \int_0^1 x_j^{1/\alpha} dj = y. \tag{4}$$

where $x_j \equiv \int_0^1 x_{i,j} di$ is the total demand for good j.

Firms choose prices and quantities to maximize their profit. Let π_j denote the profits of firm j, denominated in units of the numeraire c good. Profits are the price times aggregate amount sold, minus production cost:

$$\pi_j = p_j x_j - x_j^{1/\alpha} \tag{5}$$

The competitive firms are zero-profit firms by definition. Therefore, integrating over all x-firms in the economy yields aggregate profits: $\Pi = \int_0^1 \pi_j dj$. Each household is endowed with an equal share of all the firms in the economy. Therefore firm profits are distributed equally to households.

Equilibrium An equilibrium in this economy is

1. a set of consumption choices for each agent c_i and $x_{i,j}$ and labor input choices n_i that maximizes utility (1) subject to the budget constraint

$$c_i + \int_0^1 p_j x_{i,j} dj \leq w_i n_i + \Pi \tag{6}$$

2. a price p_j that maximizes profit (5) for each firm j, taking prices as given.

3. The markets for c goods, x goods, and labor n clear.

1.1 Results

The first order condition for labor tells us that the shadow value of wealth is $\lambda = \theta/w_i$. Substituting this into the consumption first-order condition delivers a simple relationship between consumption, productivity and the value of leisure:

$$c_i = \frac{w_i}{\theta}. \tag{7}$$

Optimal consumption of differentiated products follows a cutoff rule: Agents buy good j if the additional utility it provides exceeds the price times the shadow value of wealth: $\nu \geq p_j \theta/w_i$. Rearranging produces a cutoff rule; workers with productivity higher than the cutoff purchase of good j

$$x_{i,j} = \begin{cases} 1 & \text{if } w_i \geq \frac{\theta}{\nu} p_j \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Therefore, the fraction of agents who buy a differentiated product is the probability that each agent has a labor productivity higher than the cutoff value:

$$x_j \equiv \int_0^1 x_{i,j} di = 1 - G((\theta p_j/\nu - z)/\sigma) \tag{9}$$

where G is the cumulative distribution function associated with g.

The last piece of the consumer's problem is the labor choice. The budget constraint tells us that labor supply is

$$n_i = \frac{1}{\theta} + \int_0^1 \frac{p_j}{w_i} x_{i,j} dj - \frac{\Pi}{w_i}. \tag{10}$$

Output and GDP Since agents buy good x_j when $w_i \geq p_j \theta / \nu$. It turns out that $p_j = p$ for all goods j. Therefore, (10) tells us that aggregate labor input is

$$n = \frac{1}{\theta} + \int_0^1 \frac{\mathbb{1}_{w_i \geq p\theta/\nu} - \Pi}{w_i} di \tag{11}$$

and, using (3), aggregate output of intermediate goods is

$$y = \frac{z}{\theta} + \left(1 - G\left(\frac{\theta p}{\nu} - z\right)\right) - \Pi. \tag{12}$$

This measure of intermediate goods production is not the right quantity to compare to empirical output measures, which are the value of all final goods and services produced. Therefore, we define another measure of output that we will use for data comparison.

To discuss what the cyclical properties of our aggregate variables are, we first need to define GDP in our model.

$$GDP = \int_{i=0}^1 \left(c_i + \int_{j=0}^1 p x_{i,j} dj\right) di \tag{13}$$

Note that GDP varies, not just because of changes in production, but also because of changes in the relative price of x and c goods. This is analogous to the relative price effects on real GDP in the data.

Prices and Markups Firms set prices to maximize profit, taking the agents' behavior as given. Substitute the previous expression for demand into the profit function (5): $\pi_j = p_j(1 - G(\cdot)) - (1 - G(\cdot))^\alpha$. Differentiating this expression with respect to price p_j yields the first order condition for profit maximization characterizing the optimal price:

$$p_j = \frac{1}{\alpha}\left[1 - G\left(\frac{\theta p_j - z\nu}{\nu\sigma}\right)\right]^{(1-\alpha)/\alpha} + \frac{1 - G((\theta p_j - z\nu)/\nu\sigma)}{g((\theta p_j - z\nu)/\nu\sigma)}\left(\frac{\nu\sigma}{\theta}\right) \tag{14}$$

8

The first term is the marginal cost. The second term is the gross markup. It is proportional to the inverse of the hazard function for the productivity distribution g. It is the role of dispersion σ in this second markup term that we will focus on. Since all firms face identical distributions of incomes, they choose to set identical prices: $p_j = p$ for all j.

Proposition 1. *The gross markup is strictly increasing in earnings dispersion σ.*

See appendix A for all proofs. The gross markup is $(\nu\sigma/\theta)H((\theta p - z\nu)/\nu\sigma)^{-1}$ The first term is increasing in σ. The second term is the inverse of a hazard function (a demand elasticity) whose argument is decreasing in σ. The assumption that g is log-concave ensures that its associated hazard function is increasing in its argument, and thus decreasing in σ. This tells us that more dispersion decreases demand elasticity. When the elasticity of demand is low, firms lose fewer sales from having high markups and thus raise their markup.

Proposition 2. *The price p of x-goods is strictly increasing in earnings dispersion σ.*

When dispersion increases, the gross markup increases and the marginal cost increases. Marginal costs rise because more agents buy x goods when their earnings dispersion is higher. Not every set of assumptions would deliver this. Because our distribution of earnings shocks is truncated at zero, it is an asymmetric distribution. When dispersion rises, more agents have earnings in the right tail, but not more in the left tail. This result could also arise in settings with symmetric distributions. For example, if g was the normal distribution, which also has an increasing hazard function, an increase in earnings dispersion would increase gross markups and prices as long as $\theta p - z\nu > 0$. This condition is equivalent to requiring that half of all agents purchase the good in equilibrium.

Here, we have analyzed the gross markup because it illustrates clearly what dispersion does to the prices relative to marginal costs. In the data, researchers typically measure the percentage markup. A percentage markup (hereafter 'markup') is the percentage difference

between the price and the marginal cost of production, both denominated in y goods

$$\text{markup}_j := \log\left(\frac{\alpha p}{x_j^{(1-\alpha)/\alpha}}\right). \qquad (15)$$

Since a fraction $1-G((\theta p_j/\nu-z)/\sigma)$ of agents buys each x good, the markup can be expressed as $\exp(\text{markup}) = (\alpha\nu\sigma/\theta)(1 - G(\theta p_j/(\nu\sigma) - z/\sigma))^{(\alpha-1)/\alpha}h(\theta p_j/(\nu\sigma) - z/\sigma)^{-1}$. When we examine the quantitative implications of the model, we will use this measure of markups.
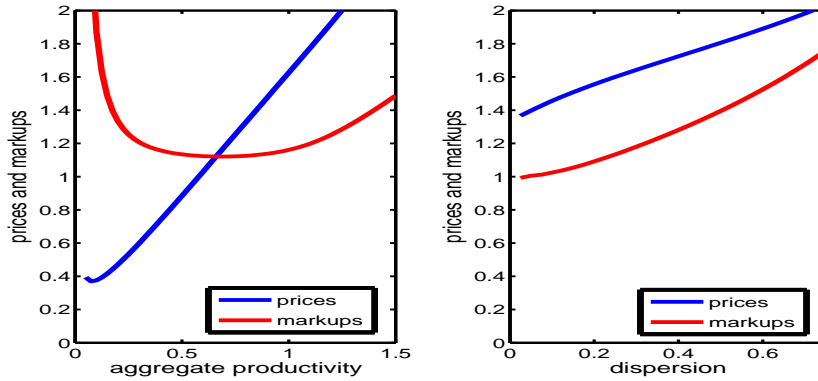

Figure 2: THE EFFECT OF PRODUCTIVITY AND DISPERSION ON PRICES AND MARKUPS.
.

Figure 2 illustrates how prices and markups react to changes in productivity and dispersion (for our calibrated parameters). The reason markups are counter-cyclical is apparent: Recessions are times when productivity is low and dispersion is high. Productivity has little effect on markups unless it is extremely high or low. Dispersion increases markups. Therefore, markups are high in recessions.

The figure also illustrates the non-linearity that makes the model difficult to calibrate but rich in predictions. When productivity is very low and rising, markups fall and then rise. If productivity low, buyers are poor and elasticity is low. Therefore, sellers follow a low-quantity high-markup strategy but earn low profits overall. As productivity rises and sales become more probable, sellers can make higher profits with lower markups that increase sales volumes. If we were in the high-productivity region (left graph, right side), the mechanism

10

could break down because the productivity effect could cause markups to rise in booms. To know know what does happen in a realistic parameter range, we turn to calibration.

2 Dynamic model

Our dynamic model is simply a repeated static model where productivity z and earnings dispersion σ fluctuate. Time is discrete and infinite $t = 0, 1, \ldots$. Productivity follows an AR(1) process:

$$\log(z_t) = (1 - \rho)\log(\bar{z}) + \rho \log(z_{t-1}) + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2). \tag{16}$$

2.1 Calibration

We begin with the dynamics of labor productivity, which we infer from data on earnings. Storesletten, Telmer, and Yaron (2004) estimate an earnings process that has idiosyncratic earnings shocks with persistent and transitory components:

$$\varepsilon_{it} = \xi_{it} + u_{it}, \qquad u_{it} \sim \mathcal{N}(0, \sigma_u^2) \tag{17}$$

$$\xi_{it} = \rho_\xi \xi_{it-1} + \eta_{it}, \qquad \eta_{it} \sim \mathcal{N}(0, \sigma_{\eta t}^2), \tag{18}$$

and dispersion that increases when productivity is below average

$$\sigma_{\eta t} = \begin{cases} \sigma_h & \text{if} \quad z_t \geq \bar{z}, \\ \sigma_l & \text{if} \quad z_t < \bar{z}. \end{cases} \tag{19}$$

These estimates have been controversial, because of the difficulty identifying transitory and permanent shocks. Guvenen (2005) and others argue that, because of unmeasured permanent

11

differences in earnings profiles, the persistence of earnings shocks is overestimated. While this distinction is crucial in a consumption-savings problem, it is not relevant for our aggregate model. Whether earnings dispersion is persistent because each person gets persistent shocks or because new workers with more dispersed characteristics enter the sample — this does not matter to our seller who sets the price and determines the probability of trade. Thus both sides in this debate hold views consistent with our model's predictions.

In the model, an agent with productivity w_i earns w_i/θ, plus an additional amount p if his productivity exceeds the threshold (8) for consuming x goods. These earnings vary across individuals both because of variation in the rate of payment per unit of time w_i and differences in hours worked (the p component). Earnings from firms profits Π are assumed to be split equally and therefore do not contribute to dispersion. The relationship between earnings and output in the model is not something we can manipulate directly because both earnings and GDP are endogenous variables. Therefore, we use the data to craft a process for exogenous variables – aggregate and idiosyncratic labor productivity – that produces endogenous series that resemble the data. To do this mapping, we make two departures from the STY framework. First, they estimate an earnings dispersion process that changes volatility, depending on whether *output* is above or below trend. Since output is endogenous, we specify that volatility depends on whether *productivity* is above or below trend. Then, we calibrate the mean of aggregate productivity to ensure the income dispersion and output have the same correlation as in the data. Second, we simulate the STY process and match its mean and standard deviation in each period to a gamma-distribution for productivity. The gamma has a lower bound of zero, ensuring that productivity is never negative and has the log-concave property that we used for the theoretical results. In practice, a gamma distribution looks very similar to a log normal frequently used to model wages and earnings. Appendix B details all our data and its transformations.

utility weight on x goods	ν	2.2	average markup	0.2
utility weight on leisure	θ	1.65	average hours	0.3
concavity of production	α	0.6	profit share π/GDP	0.33
mean of productivity	$\bar{z}$	1	corr(σ, GDP)	-0.3
productivity innovation s.d	σ_z	0.0065	output s.d.	0.02
productivity autocorrelation	ρ	0.85	output autocorrelation	0.8
transitory earnings volatility	σ_u	0.065	STY estimate	0.065
persistent earnings volatility $z > \bar{z}$	σ_h	0.032	STY estimate	0.032
persistent earnings volatility $z < \bar{z}$	σ_l	0.054	STY estimate	0.054
earnings autocorrelation	ρ_ξ	0.988	STY estimate	0.988

Table 1: Parameter values for the simulated model and the features of the data each parameter matches.

Table 1 contains a summary of the calibrated parameters. Appendix B details the precise transformations of the data and the model that deliver these targets. Standard calibration targets are 1/3 of time spent on labor and a profit share of 2/3 to pin down θ and α. We calibrate the productivity process z_t to match the persistence and standard deviation of output as reported in Stock and Watson (1999). What is not standard about the calibration is that we do not normalize the mean $\bar{z}$ of the productivity process to 1. Instead, we calibrate it to the correlation of dispersion with output. Calibrating $\bar{z}$ is required because the model is not scale-neutral. Matching this correlation ensures that our dispersion process has the same properties as the one Storesletten et. al. estimate.

Markups are defined as in (15). Using manufacturing data, Basu (1996) argues that markups are small in the aggregate economy, in the range of 10% or less. Using data on automobiles and cereal, Berry, Levinsohn, and Pakes (1995) and Nevo (2001) document markups of 27-45%. This points to large heterogeneity in markups across sectors. Our markup refers to the x good sector only. In the c sector, markups are 0%. Averaging the two sectors delivers an 11% expenditure-weighted average markup. The expenditure weight depends on the probability of a a randomly-chosen agent buying x goods. That unconditional average probability of trade is 77%.
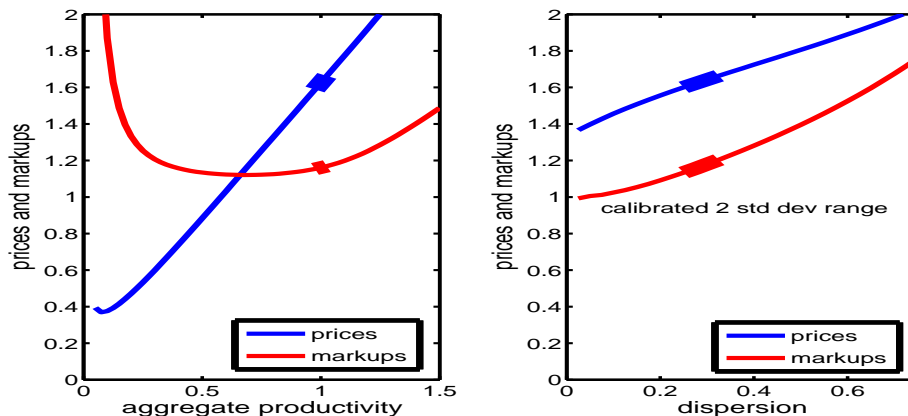
Figure 3: CALIBRATED RANGE OF PARAMETERS.
In the left panel, the shaded part of the line is the 2 standard deviation range of productivity, . In the right panel, the 2 standard deviation range of earnings dispersion is shaded. Both are based on the parameter estimates in table 1.

One of the reasons that it is important to calibrate the model is because the cyclical behavior of prices cannot be inferred from theory alone. Only when the relevant region of the parameter space is identified can the model's predictions be compared to data. The left hand panel of figure 3 is the same theory-based mapping between productivity and price as figure 2, with the calibrated 2 standard deviation interval of productivity around its mean highlighted. In this interval, prices are increasing in productivity and dispersion. Since productivity and dispersion move in opposite directions over the business cycle, a calibrated model is required to predict the cyclical movement of prices.

2.2 Markups

We characterized recessions as times when firms pursue low-volume, high-margin sales strategies. To measure this effect in our model, we examine the cyclical properties of markups, as defined in (15). The correlation of markups and log output in the simulated model is -0.09. The standard deviation of markups is 0.65 times the standard deviation of log output. In contrast, in a perfectly competitive market, the markup would always be zero. Figure 4

14

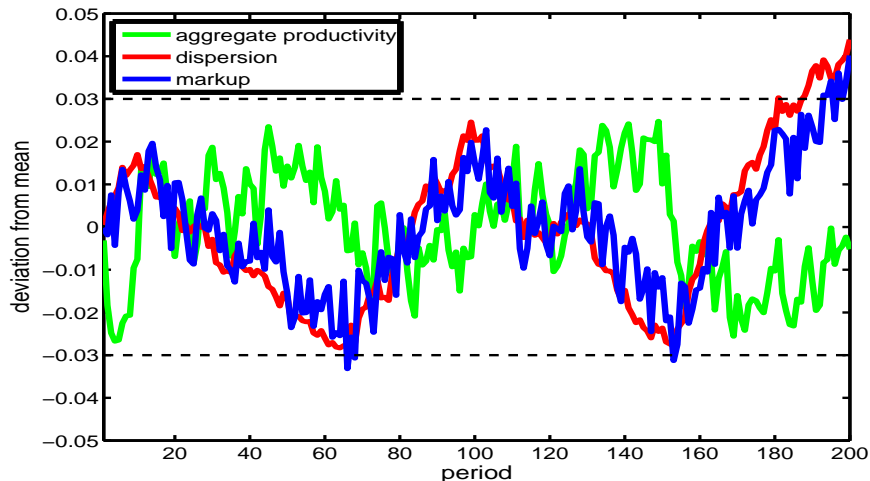illustrates a random simulated time-series of markups.

Figure 4: MARKUPS, EARNINGS DISPERSION AND PRODUCTIVITY IN THE SIMULATED MODEL.
Based on calibrated parameters in table 1.

In the data, counter-cyclical markups have been documented by and Rotemberg and Woodford (1999) using three different methods, by Murphy, Shleifer, and Vishny (1989) using input and output prices, by Chevalier, Kashyap, and Rossi (2003) with supermarket data, by Portier (1995) with French data and by Bils (1987) inferring firms' marginal costs.

Besides their negative correlation with output, the other salient cyclical feature of markups is that they lag output. Table 2.2 shows that the model's markup is positively correlated as a leading variable and mostly strongly, negatively correlated as a lagging variable, just as in the data. The difference is that the model's markup is only half as counter-cyclical as in the data.

	-4	-3	-2	-1	0	1	2	3	4
Data	-0.30	-0.34	-0.34	-0.31	-0.21	-0.05	0.08	0.19	0.29
Model	-0.31	-0.26	-0.19	-0.11	-0.09	0.02	0.05	0.07	0.08

Table 2: LEADS AND LAGS OF MARKUP-GDP CORRELATIONS.

Entries are corr($\ln(\text{markup}_t)$, $\ln(\text{GDP}_{t+k})$). Positive numbers indicate leads and negative numbers indicate lags. Data from Rotemberg and Woodford (1999) (table 2, column 2). Markup is estimated using the labor share in the nonfinancial corporate business sector and an elasticity of non-overhead labor of -0.4.

The reason that the model's markups are lagging is that the earnings dispersion process is highly persistent. In low-productivity periods, the shocks to individual's earnings start to become more volatile. Specifically, the shocks to the persistent component of earnings grow in magnitude (equation 18). As these high-volatility shocks continue to arrive, the earnings distribution fans out. When productivity picks up and shocks become less volatile, there is enormous dispersion in the persistent component of earnings that takes a long time to revert to its mean. It takes many periods of low-volatility shocks for the earnings distribution to become less dispersed. Since markups are driven by earnings dispersion, which is a lagging variable, markups are lagging as well. This feature of the model is similar to Bilbiie, Ghironi, and Melitz (2006) who generate lagged markups because firm entry lags productivity and markups depend on the number of firms. In their setting, it is the large fixed cost of firm entry that causes firms to delay their entry decisions until productivity is sufficiently high.

2.3   Predictions for Other Macroeconomic Aggregates

The goal of the paper is to explain markups and, at the same time, match business cycle quantities. Therefore, we compare three macroeconomic aggregates to the GDP measure defined in (13): the share of output that firms earn as profits (π/GDP), labor inputs n and the real wage. The real wage is a measure of relative prices. Comparing prices of x and c goods to a measure like the CPI has the problem that the CPI is the rate of exchange between goods and money. Yet there is no money in this model. Therefore, we report a relative price we can interpret: the relative price of labor to the expenditure-weighted price index of x and c goods.

Table 3 compares the model aggregates to data. Profit shares have cyclical properties that do a reasonable job of matching the data. Most importantly, profits are procyclical.

model variable	std dev.	corr with GDP
profit share (π/GDP)	0.69	0.20
real wages	0.38	0.57
labor	2.38	0.12
data variable	std dev.	corr with GDP
profit share (π/GDP)	0.80	0.22 (0.37)
real wages	0.39 (0.36)	0.16 (0.25)*
labor (employment)	0.84 (0.82)	0.81 (0.89)
labor (hours)	0.97 (0.98)	0.88 (0.92)

Table 3: Second moments of aggregate variables in the model and data.
Standard deviations are divided by the standard deviation of GDP. Most statistics are from Stock and Watson (1999). Labor and wage numbers in parentheses are from Cooley and Prescott (1995). Number with an asterisk is from Rotemberg and Woodford (1999) All capital share statistics come from the labor share statistics reported in Gomme and Greenwood (1995). The second correlation, in parentheses, comes from NIPA data. But the NIPA counts all proprietors' earnings as profits, although it is part profit and part labor compensation. The first correlation corrects for this by removing proprietor's earnings.

This is an important piece of evidence that distinguishes this model from sticky price theories or models with free-entry. Profits are only slightly less volatile in the model as in the data.
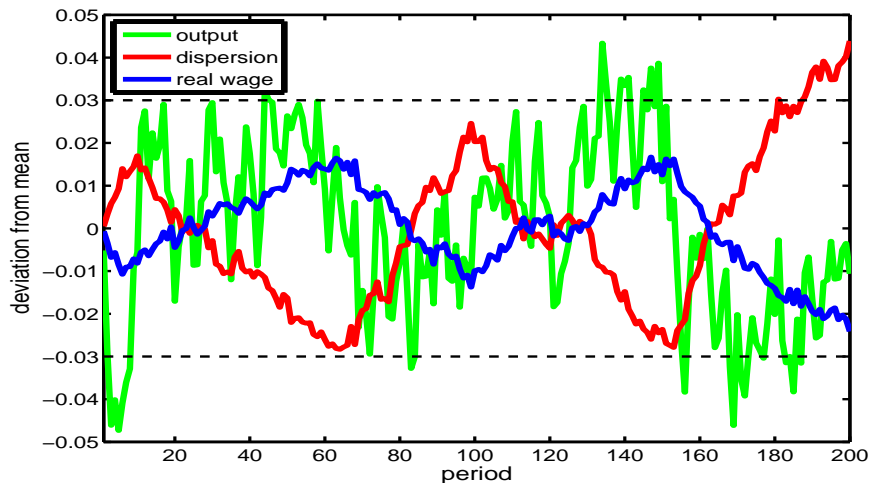

Figure 5: REAL WAGES, EARNINGS DISPERSION AND GDP IN THE SIMULATED MODEL. Based on calibrated parameters in table 1.

Real wages do slightly less well; they are too cyclical, but have the right volatility. Without a capital stock in the model, wages and output are more driven by changes in productivity, than they would be with capital. This makes the wage-output correlation high. Real wages are also a useful measure of price movements because many people use its

inverse as a measure of markups. The fact that simulated real wages are pro-cyclical means that this alternative measure of the model's markups is counter-cyclical as well.

Figure 5 illustrates the behavior of real wages and GDP. It has two features that look familiar. First, real wages look like the mirror image of markups. The intuition for this is that wages are the main component of marginal costs and so wages relative to p behaves like the reciprocal of the markup. Furthermore, both real wages and markups are closely correlated with dispersion. Second, the measure of GDP looks quite similar to the productivity plotted in figure 4. This tells us that, although dispersion has an effect on GDP, it is still primarily driven by productivity shocks. Time-varying earnings dispersion increases the volatility and persistence of GDP. This can be seen in the lower than normal calibrated standard deviation and autocorrelation of the productivity shocks used to match the second moments of output. Instead of he typical 0.0072 standard deviation and 0.98 persistence used to match the properties of Solow residuals, ours are lower: 0.0065 standard deviation and 0.85 autocorrelation.

The labor predictions are the biggest failure of the model. Labor in the model looks unlike the data. The reason why labor is insufficiently cyclical comes from a wealth effect: When productivity rises, labor should increase. But, at the same time, firm profits, which get rebated to workers, increase as well. This increase in earnings for workers makes them value additional earnings less, relative to additional leisure and induces them to work less. While the model fails to reproduce the correlation of labor and output, this is a failing of many models. Furthermore, it has a well-known solution that can potentially be implemented in our framework. Appendix C introduces a modification to preferences to correct this problem.

2.4 Benchmark Economies

To better understand what drives our results, we compare our model to two benchmarks. The first is an economy where earnings dispersion is constant, always equal to its mean value in the full model. The second benchmark is an economy where there is no earnings dispersion, only a representative consumer.

When dispersion is constant and equal to its unconditional mean, many of our calibration targets look similar. The average markup (16%), the average labor supply (0.28), average profit share (0.33), the standard deviation of output (0.02) and autocorrelation of output (0.8) are all essentially unchanged. The one big difference is that the correlation of profit share and output (0.76) is much higher than its previous value (0.2).

But this benchmark model fails to deliver counter-cyclical markups or realistic real wages. The correlation of markups and GDP is 0.33, meaning that markups switched from being counter-cyclical to pro-cyclical. Their standard deviation relative to GDP is 0.31. For real wages, they are far too correlated with GDP (0.99, versus 0.16 in the data) and not volatile enough (standard deviation relative to GDP is 0.13). In addition, profit shares are not sufficiently volatile relative to output (std=0.39), and the benchmark model shares all the shortcomings of the full model with respect to its labor predictions (std and correlation relative to GDP are 2.3 and 0.23).

In the second benchmark, when earnings dispersion is zero, aggregates are either insufficiently volatile or almost perfectly cyclical. The main problem is that agents all work, all the time $n_{i,t} = 1$. So, the correlation and standard deviation of labor relative to output is zero. Because of this, output is not sufficiently volatile (0.012), the average labor payments are too high, making average profit shares too low (0.17). Because payments to labor just vary with productivity (correlation 0.94), the firms' profit share is also almost perfectly correlated with output (0.99) and far too volatile (std 3.05). Defining markups in this setting is not

obvious because sellers either sell 1 or 0 units, making marginal cost sometimes zero, and markups undefined. Defining the markup as price/average cost, we find that the average markup is a somewhat high 33%. However, that markup is perfectly cyclical and as volatile as GDP. This stark contrast makes clear that the effects of our model truly are being driven by the earnings dispersion mechanism and that both the presence of earnings dispersion and its time-variation are essential features of the model.

2.5 Market Inefficiency and the Cost of Business Cycles

One of the reasons that researchers have spent so much time trying to understand the source of counter-cyclical markups is that they produce prices that vary little over the cycle. In the business cycle literature, such price rigidity is important because it makes business cycles more costly than they are in models where prices vary more. Lucas (1987) estimates that the aggregate welfare cost of business cycles in a standard model is a meager 0.008-0.1% of consumption. Yet, enormous resources are devoted to managing business cycles and many people believe them to be an important source of risk. Counter-cyclical earnings dispersion can help to reconcile these two points of view.

Adding time-varying heterogeneity in productivity amplifies the cost of business cycles for some agents because it increases markups and reduces the size of the x sector in downturns. This means that the fraction of the population who are hit by bad productivity shocks are forced to consume a more limited variety of goods. This amplifies the welfare effect of the shock for those unlucky agents. In our simulated model, the cost of the business cycle for the median agent is small, 0.27%. This is still three times larger than the Lucas estimate. However, for the agents in the 10th and 1st percentiles of the earnings distribution, they would be willing to sacrifice 1.8% and 3.6% of their consumption to eliminate aggregate

fluctuations.[2] These are large welfare costs.

3 Long-Run Implications of the Model

3.1 The long-run decline in business cycle volatility

While the model is built to explain fluctuations as business cycle frequencies, there has been a long-run increase in the level of earnings dispersion that should cause low-frequency changes as well. One of the most discussed low-frequency changes in the U.S. economy has been the decline of business cycle volatility. The observed increases in productivity and in earnings dispersion generate a comparable decline in GDP volatility in our model.

We simulate 5 model economies, each representing a decade from the 50's to the 90's. To make the time-averaged moments of our growing economy like the moments in our benchmark business cycle model, we choose the middle decade (the 70's) to be the same as our benchmark calibration. Then, we match the decade-by-decade growth in productivity and earnings dispersion, using data on labor productivity from the BLS and earnings dispersion from Heathcote, Storesletten, and Violante (2006). (See appendix B for data details.) We simulate each decade as if it were a separate economy and report its GDP volatility in table 4.

According to Heathcote, Storesletten, and Violante (2006), earnings dispersion increased by 20% from 1967-1996, an average annual rate of 0.66%. In the model, as earnings dispersion grows, fewer agents are close to marginal and aggregate demand becomes less elastic. Shocks to labor productivity have less effect on who buys what products. Since producers are producing in anticipation of changes in aggregate demand, when aggregate demand becomes

---

[2]These welfare costs are long-run averages. An agent who is in the 10th percentile of the distribution during a boom does not have the same welfare cost as an agent who is in the 10th percentile during a recession. Because agents who are currently at a given place in the earnings distribution will not occupy the same position quarters later, the exact welfare cost depends on the point in the business cycle when the question is posed.

decade	model	data
1950's	2.0%	2.1%
1960's	2.1%	1.3%
1970's	1.6%	2.2%
1980's	1.4%	1.7%
1990's	1.3%	0.9%

Table 4: BUSINESS CYCLE VOLATILITY, STD(LN(GDP)), AS PRODUCTIVITY AND EARN-INGS DISPERSION INCREASE.
Data are standard deviations of quarterly GDP, by decade.

less volatile, GDP volatility falls as well. The 20% rise in earnings dispersion results in a 40% drop in the standard deviation of GDP.

3.2 Long-run slowdown in real wage growth

A second long-run change that has been of particular concern to policy-makers is the slow-down in the growth of real wages. In figure 6, the left panel illustrates how real wages were keeping pace with productivity growth until the 1970's, when real wage growth slowed down.
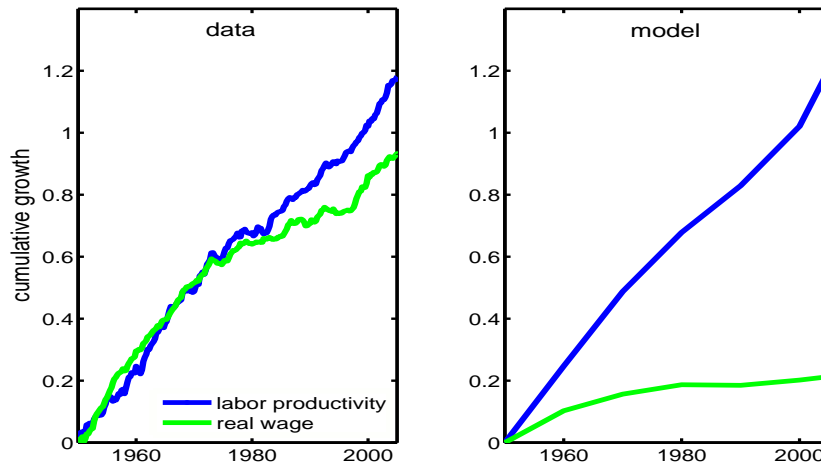

Figure 6: PRODUCTIVITY AND WAGE GROWTH IN THE DATA AND THE SIMULATED MODEL. The trend added to the model is the decade-by-decade increase in in labor productivity and in earnings dispersion estimated by Heathcote, Storesletten, and Violante (2006). All other parameter values are listed in table 1. Data: See appendix B for details.

We ask if feeding in the increase in productivity and the increase in earnings dispersion

can generate this real wage time series in our model, using the same exercise as in the previous section. We simulate each decade as if it were a separate economy and report its average productivity and wages in the right panel of figure 6.

When productivity and earnings dispersion rise, markups rise and the share of output paid to capital rises. Because firms are exploiting the decrease in demand elasticity, prices rise faster than wages.

The flip side of this finding is that firms' profit shares are rising. In the model, the share of output rises steadily from 25% in 1950's to 33% (our calibrated value) in the 1970's to 48% in the 2000's. The evidence of the size of this rise in the data is mixed. The share of output not paid out as labor earnings – a very broad definition of profits – rose only by about 5% from 1970-96 (NIPA data). Meanwhile, net payouts to security holders as a fraction of each firm's value added – a much more narrow definition of profits – rose from 1.4% to 9% (flow of funds data) or 2.3% to 7.5% (NIPA data). While the broad measure suggests that our model over-predicts the rise in profits, the latter measure suggests that it is a much larger phenomenon than what the model can replicate.

4   Cross-Sectional Implications of the Model

We have used our model to argue that time-variation in earnings dispersion should cause prices to change at business cycle and long-run frequencies. The same mechanism should also operate in the cross-section. States or social groups with more earnings heterogeneity also have lower demand elasticity. This section examines two pieces of evidence that suggest cross-section price patterns are consistent with our theory.

4.1 Testing the model with state-level panel data

The model's key mechanism is that higher earnings dispersion induces sellers to raise prices. Using panel data for 49 U.S. states, we find that earnings dispersion and prices do have a significant positive relationship. Furthermore, the model predicts that productivity raises prices, but that productivity moderates the effect of earnings dispersion. Both of these effects are supported by the data as well.

We begin by estimating OLS regression coefficients in our model to derive our predictions. Three predictions emerge. First, higher dispersion raises price because more dispersion reduces the price elasticity of demand. That induces sellers to raise prices and the model CPI rises. Second, higher productivity makes the average buyer have a higher willingness to pay. In contrast to a representative agent business cycle model where the relative price of a good produced with a productive technology is low, our relative x-good prices are higher when those goods can be produced more cheaply. These two effects can be seen in figure 3.

The third effect is a more subtle interaction between productivity and dispersion. When aggregate productivity increases, firms raise prices because buyers have more ability to pay. The fact that firms make more profits on each unit also makes them more averse to the possibility of losing the sale. Therefore, they don't increase prices one-for-one with the change in productivity. When income dispersion is small, sellers can sell to almost the entire distribution. But raising their optimal price by a little bit loses many customers. When that low-dispersion distribution shifts, the optimal prices moves with it almost one-for-one. When income dispersion is higher, firms are making fewer sales and larger profits per sale. In these situations, firms react to increases in productivity by raising prices only slightly and enjoying an increase in market share. This interaction effect is small. The standard deviation of the probability of trade is 1.3%. A 1% change in this probability reduces the effect of a 1% change in dispersion by about 0.05%.

Next we do the same estimation on our state-level data. To measure a state's earnings dispersion, we take the log average salaries in each of the state's counties, weight them by the number of jobs in the county, and take their cross-sectional standard deviation. State price levels come from Del Negro (1998). As a proxy for state productivity, we use real state GDP per employed worker. Data sources and the details of data manipulation are in appendix B.

	Income dispersion (σ)	Productivity ($\log(z)$)	Interaction ($\sigma \log(z)$)
Model	0.21* (0.045)		
	0.41* (0.001)	0.79* (0.001)	
	0.41* (0.001)	0.79* (0.026)	-0.05 (0.091)
Data	0.11* (0.027)		
	0.09* (0.028)	0.04* (0.009)	
	0.06* (0.034)	0.04* (0.009)	-0.20 (0.147)

Table 5: EFFECTS OF DISPERSION AND PRODUCTIVITY ON PRICE, IN STATE-LEVEL PANEL DATA. Coefficients are OLS estimates of: $\text{price}_{s,t} = \text{constant}_s + \beta_1 \sigma_{s,t} + \beta_2 \log(z_{s,t}) + \beta_3 \sigma_{s,t} \log(z_{s,t}) + \varepsilon_{s,t}$. For both model and data, row 1 constrains $\beta_2 = \beta_3 = 0$. Row 2 constrains $\beta_3 = 0$. Row 3 is the unrestricted estimates. Standard errors in parentheses. A * denotes coefficients significant at the 1% level.

Table 5 shows that all three predictions are upheld, qualitatively, in the data. However, a few caveats are in order. First, the size of the effects are quite different in the model from what they are in the data. Of course a model with no additional source of shocks should produce a somewhat stronger statistical relationship than the data. Second, the fraction of variation in the data's prices that are explained by productivity and earnings dispersion are small. The R^2s of the regressions are all below 5%. Finally, this exercise does not rule out the possibility that some external factor is causing both prices and earnings dispersion to vary. However, we did remove any nationwide source of variation. Prices, productivity and earnings dispersion are all measured as deviations from their national averages. Thus if some external force is generating these covariances, it must be an effect orthogonal to aggregate fluctuations.

4.2   Support from Empirical Pricing Studies

Our results are also qualitatively consistent with the findings of Chevalier, Kashyap, and Rossi (2003). Periods of good-specific high demand (e.g., beer on the fourth of July) are times when consumers' values for the goods are more similar. While one might expect that high demand would increase prices, the authors find that prices and markups fall. The same outcome would arise in our calibrated model if productivity dispersion σ falls.

More support for our mechanism comes from a study of the effect willingness-to-pay dispersion has on car sales. Goldberg (1996) estimates that blacks' valuations for new cars are more dispersed than whites'. She then collects data on the initial offer to blacks and whites by a car salesman. The initial offer price is higher, and the probability of sale lower, for the group with more dispersed willingness to pay.

5   Conclusion

Our production economy is set up to capture the simple intuition that when earnings dispersion is higher, the price elasticity of demand is lower, so sellers optimally raise markups and thus prices. However, without quantifying the model, the cyclical behavior of prices and markups is ambiguous because the relative size of the productivity and earnings dispersion effects cannot be evaluated. Using estimates of the time-series variation in the earnings distribution, we calibrate the model. Although the model is a simple one, it does a reasonable job of matching many features of prices and traditional business cycle quantities. Where the model fails to replicate features of the data, it is because of well-known problems with well-known fixes. These results suggest that new research measuring the relationship between aggregate output and individual-level heterogeneity can help to reconcile quantity-oriented business cycle theory with monetary facts about counter-cyclical markups. If so, this would

be an important step forward. Explaining counter-cyclical markups can help us understand why prices appear rigid, varying less over the cycle than competitive models predict. Rigid prices make business cycles more costly but also create a role for monetary policy to stabilize business cycle fluctuations.

A   Proofs

A.1   Proof of proposition 2

From the first-order condition for prices (14), define a function F such that when $F = 0$, the first-order condition holds:

$$F \equiv \frac{1}{\alpha}(1 - G\left(\frac{\theta p_j - z\nu}{\nu\sigma}\right))^{(1-\alpha)/\alpha} + \frac{1 - G((\theta p_j - z\nu)/\nu\sigma)}{g((\theta p_j - z\nu)/\nu\sigma)}\left(\frac{\nu\sigma}{\theta}\right) - p_j.$$

Define $\psi \equiv (\theta p_j - z\nu)/(\nu\sigma)$ and differentiate F with respect to prices and dispersion.

$$\frac{\partial F}{\partial p} = -\frac{1-\alpha}{\alpha^2}(1 - G(\psi))^{(1-2\alpha)/\alpha}g(\psi)\frac{\theta}{\nu\sigma} - H(\psi)^{-2}H'(\psi)$$

Since $(1 - G(\cdot)) \geq 0$ and $g(\cdot) \geq 0$ for any cumulative and probability density functions G and g, the parameters θ, ν are assumed to be positive, σ is a standard deviation, and thus must be non-negative and $\alpha < 1$, the first term is negative, due to the minus sign in front. In the second term, any hazard function is positive $H \geq 0$. Since we assumed our distribution is log-concave, $H'(\cdot) \geq 0$ as well. Therefore, $\partial F/\partial p \leq 0$. Next, differentiate F with respect to dispersion.

$$\frac{\partial F}{\partial \sigma} = \frac{1-\alpha}{\alpha^2}(1 - G\left(\psi\right))^{(1-2\alpha)/\alpha}g(\psi)\frac{\psi}{\sigma}$$

$$+H(\psi)^{-2}H'(\psi)\left(\frac{\psi}{\sigma}\right)\left(\frac{\nu\sigma}{\theta}\right) + H(\psi)^{-1}\left(\frac{\nu}{\theta}\right)$$

Rearranging and substituting the first partial derivative into the second yields

$$\frac{\partial F}{\partial \sigma} = H(\psi)^{-1}\left(\frac{\nu}{\theta}\right) - \left(\frac{\nu\psi}{\theta}\right)\frac{\partial F}{\partial p}.$$

Since the hazard function H and the parameters ν, θ are non-negative, the first term is positive. Suppose that $\psi \geq 0$. Then, since $\partial F/\partial p \leq 0$, but enters with a negative coefficient, the second term would be positive as well. Therefore, if $\psi \geq 0$, $\partial F/\partial \sigma \geq 0$.

Show that $\psi \geq 0$. Since $\nu, \sigma \geq 0$, this is equivalent to showing that $\theta p_j \geq z\nu$. Suppose not. Substituting $p_j < z\nu/\theta$ into the first order condition (14), delivers a second term on the right side that takes the form $(1 - G(\psi)/g(\psi))\nu\sigma/\theta$ for $\psi < 0$. Recall the assumption that the distribution g puts zero probability on negative entries. Thus, $g(\psi) = 0$ which makes the second term infinite.

Since p_j and the first term are finite, the first-order condition cannot hold. Therefore, any solution to the optimal price problem involves $\psi \geq 0$.

Finally, apply the implicit function theorem. It tells that that if $\partial F / \partial p$ and $\partial F / \partial \sigma$ have opposite signs, $\partial p / \partial \sigma > 0$. Since we have already shown $\partial F / \partial \sigma \geq 0$ and $\partial F / \partial p \leq 0$, this completes the proof.

A.2 Proof of proposition 1

Define a function F_1, which is equal to zero when the optimal gross markup m_j is chosen:

$$F_1 \equiv \frac{1 - G((\theta\chi + \theta m_j - z\nu)/\nu\sigma)}{g((\theta\chi + \theta m_j - z\nu)/\nu\sigma)} \left(\frac{\nu\sigma}{\theta}\right) - m_j$$

where χ is the marginal cost of production.

As before, define $\psi \equiv (\theta(\chi + m_j) - z\nu)/(\nu\sigma)$ and differentiate F_1 with respect to markups and dispersion.

$$\frac{\partial F_1}{\partial m} = -H(\psi)^{-2} H'(\psi)$$

For any hazard function, $H \geq$. Since we assumed our distribution is log-concave, $H'(\cdot) \geq 0$ as well. Therefore, $\partial F_1 / \partial m \leq 0$.

Next, differentiate F_1 with respect to dispersion.

$$\frac{\partial F_1}{\partial \sigma} = H(\psi)^{-2} H'(\psi) \left(\frac{\psi}{\sigma}\right) \left(\frac{\nu\sigma}{\theta}\right) + H(\psi)^{-1} \left(\frac{\nu}{\theta}\right)$$

$$= H(\psi)^{-1} \left(\frac{\nu}{\theta}\right) - \frac{\partial F_1}{\partial m} \left(\frac{\psi\nu}{\theta}\right)$$

Since the parameters are positive, the hazard function is positive and $\partial F_1 / \partial m \leq 0$, we know that $\frac{\partial F_1}{\partial \sigma} \geq 0$. Applying the implicit function theorem delivers the result.

B Calibration and data details

Income dispersion The quarterly persistence and standard deviation of income are derived from the annual estimates of Storesletten, Telmer, and Yaron (2004) as follows: $\varphi = 0.952^{1/4}$, the standard deviation to the persistent component is $0.125Q$ when productivity is above average and is $0.211Q$ when productivity is below average while the standard deviation of the transitory component is $0.255Q$ where the adjustment factor is $Q := 1/(1 + \varphi + \varphi^2 + \varphi^3) = 0.2546$. We deviate from Storesletten, Telmer, and Yaron (2004) in two ways. First, our shocks are gamma-distributed not normal. Second, our shock changes variance, depending on whether aggregate productivity, not output, is above its mean or not. Because of the high correlation of productivity and output, these two processes are virtually indistinguishable. We make this change because conditioning the variance on endogenous output would require solving a fixed point problem, in every period, with no guarantee of a unique solution.

To map the STY process into the model's distribution of labor productivity, we use the following procedure.

1. Simulate an AR1 for log aggregate productivity z with normal errors.

2. For each period t, determine if aggregate productivity is above average (or not) to work out which innovation standard deviation to use for the persistent component of idiosyncratic shocks e.

3. For each individual i and in each time period t, draw innovations to the persistent component of earnings, independently, from a normal distribution with the standard deviation determined in step 2. Add on the transient component also drawn from a normal distribution.

4. Calculate the cross sectional standard deviation of earnings for each period.

5. Select two parameters that govern the mean and standard deviation of the gamma distribution for each period. Set the standard deviation of labor productivity equal to the standard deviation calculated in the previous step. Set the mean of labor productivity equal to the level of the aggregate productivity from step 1.

6. Given the distribution of productivity, solve the seller's pricing problem and calculate other endogenous variables.

Simulations All simulations in this paper begin by sampling the exogenous state variables for a 'burn-in' of 1000 quarters. This eliminates any dependence on arbitrary initial conditions. A cross-section of 2500 individuals is tracked for 200 quarters, corresponding to the dimensions in Storesletten, Telmer, and Yaron (2004). Realizations of endogenous variables are then computed. The moments discussed in the text are averages over the results from 100 runs of the simulation (that is, averages over $200 \times 100 = 20000$ observations).

Aggregate data All data is quarterly 1947:1-2006:4 and seasonally adjusted. Real GDP is from the Bureau of Economic Analysis (BEA). This is nominal GDP deflated deflated by the BEA's chain-type price index with a base year of 2000. We measure aggregate labor productivity and real wages by, respectively, real output per hour and real compensation per hour in the nonfarm business sector, both from the Bureau of Labor Statistics (BLS) Current Employment Survey (CES) program. Nominal output and compensation are deflated by the BLS's business sector implicit price deflator with a base year of 1992.

State-level data We construct a balanced panel of income dispersion, price, and labor productivity data. The state income dispersion data come from the County Wage and Salary Summary (CA34), produced by the BEA's Regional Economic Accounts (REA). The data are reported annually from 1963-1997. Due to missing data for Alaska, we use 49 states. The District of Columbia is excluded because computing dispersion is impossible with only one county. Annual state employment and state GDP are also from the REA. The REA reports nominal state GDP data from 1963-1997 based on the Standard Industrial Classification (SIC) codes and from 1997-2005 based on the revised North American Industry Classification System (NAICS). The BEA reports state price indices, but only from 1990-2005. Therefore, to construct real state GDP numbers we use state consumer price indices constructed by Del Negro (1998). These are annual 1969-1995. This same price data is what we use as the dependent variable in our regressions. Therefore, we have

a balanced panel of 49 states and 27 time periods for a total of 1323 observations. Since the underlying methodologies are so different, we do not attempt to splice GDP numbers from the SIC and NAICS accounts. We measure state labor productivity as real state GDP divided by total employment in that state.

We then remove the national trends in all variables. The national trend in each variable v is the average across states, with each state weighted by its total number of employed workers: $v_{national,t} = \sum_{s=1}^{49} v_{s,t} l_{s,t} / (\sum_{s=1}^{49} l_{s,t})$, where l is the number of employed workers in state s and year t. For GDP and employment, the de-trending is done by computing the log deviation of the state series from the national average: $\log(v_{s,t}) - \log(v_{national,t})$. Since dispersion is already measured in logs, de-trended dispersion is the raw difference of state dispersion and national average dispersion: $v_{s,t} - v_{national,t}$.

C   A Model Without Wealth Effects on Labor

The wealth effect on labor can be seen in equation (11), where profits Π enter negatively in the labor supply. Although more standard models often encounter this problem, it is particularly acute here because imperfect competition in x goods makes profits larger and more volatile. Greenwood, Hercowitz, and Huffman (1988) developed "GHH" preferences that eliminate the wealth effect on labor supply. These preferences improve the model's predictions for labor. However, we choose to analyze them here, rather than in the main model because they obscure the model's mechanics. In particular, the basic model is still useful because it makes clear that the main results, those for markups, do not depend on GHH preferences.

There is a continuum of agents indexed by $i \in [0,1]$, with identical preferences over a numeraire good c_i, labor n_i, and a continuum of differentiated products x_j

$$U = \ln\left(c_i - \left(\frac{\theta n^{1+\gamma}}{1+\gamma}\right)\right) + \nu \int_0^1 x_{i,j} dj. \tag{20}$$

All of the other features of the model are unchanged. The equilibrium definition is identical, with one exception. Agents maximize the utility in (20) instead of (1).

Equilibrium relationships   The first-order condition for labor choice tells us that labor depends only on the wage and on preference parameters:

$$n_i = \left(\frac{w_i}{\theta}\right)^{1/\gamma}. \tag{21}$$

This simple relationship, devoid of any wealth effect is what GHH preferences are designed to deliver. The first-order condition for x_j reveals that demand for x goods follows a cutoff rule. Agent i buys good j if

$$p_j \leq \frac{\nu}{\lambda_i} \tag{22}$$

where λ_i is agent i's shadow value of wealth. That shadow value is

$$\lambda_i = \left\{ c_i - \frac{\theta}{1+\gamma} \left(\frac{w_i}{\theta}\right)^{(1+\gamma)/\gamma} \right\}^{-1} \tag{23}$$

From the budget constraint, c goods are

$$c_i = w_i \left(\frac{w_i}{\theta}\right)^{1/\gamma} + \pi - \int_{j=0}^{1} p_j \mathbb{1}_{p_j \leq \nu/\lambda_i} dj. \tag{24}$$

The per-capita firm profit π is firm revenues minus costs,

$$\pi_j = p_j x_j - (x_j)^{1/\alpha}. \tag{25}$$

where $x_j = \int_{i=0}^{1} x_{ij} di$. Firms prices are chosen to maximize this profit, taking agents' demand functions as given. That price solves

$$p_j = \frac{1}{\alpha} x_j^{(1-\alpha)/\alpha} - \left(\frac{\partial x_j}{\partial p_j}\right)^{-1} x_j. \tag{26}$$

This system of equations is not nearly as simple as before because the cutoff rule for whether an agent buys and x good or not is no longer a simple linear relationship with their wage. However, the fixed point of this system can be solved for numerically. The next step is to solve the model numerically, simulate it with the calibrated parameters, and see whether its predictions for wages are more in line with the data.

References

ALTIG, D., L. CHRISTIANO, M. EICHENBAUM, AND J. LINDE (2004): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," Working Paper.

BASU, S. (1996): "Procyclical Productivity, Increasing Returns or Capital Utilization?," *Quarterly Journal of Economics*, 111, 719–51.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

BILBIIE, F., F. GHIRONI, AND M. J. MELITZ (2006): "Endogenous Entry, Product Variety, and Business Cycles," Working Paper.

BILS, M. (1987): "The Cyclical Behavior of Marginal Cost and Price," *American Economic Review*, 77(5), 838–855.

CABALLERO, R., AND M. HAMMOUR (1994): "The Cleansing Effect of Recessions," *American Economic Review*, 84, 1350–68.

CHEVALIER, J., A. KASHYAP, AND P. ROSSI (2003): "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 93, 15–37.

COMIN, D., AND M. GERTLER (2006): "Medium term business cycles," *American Economic Review*, 96(3), 523–551.

COOLEY, T., R. MARIMON, AND V. QUADRINI (2004): "Aggregate Consequences of Limited Contract Enforceability," *Journal of Political Economy*, 112, 817–847.

COOLEY, T., AND E. PRESCOTT (1995): "Economic Growth and Business Cycles," in *Frontiers in Business Cycle Research*, ed. by T. Cooley. Princeton University Press.

DEL NEGRO, M. (1998): "Aggregate Risk Sharing Across US States and Across European Countries," FRB Atlanta Working Paper.

GALI, J. (1994): "Monopolistic Competition, Business Cycles, and the Composition of Aggregate Demand," *Journal of Economic Theory*, 63, 73–96.

GOLDBERG, P. (1996): "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, 104(3), 622–654.

GOMME, P., AND J. GREENWOOD (1995): "On the Cyclical Allocation of Risk," *Journal of Economic Dynamics and Control*, 19, 19–124.

GREENWOOD, J., Z. HERCOWITZ, AND G. HUFFMAN (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78(3), 402–417.

GUVENEN, F. (2005): "An Empirical Investigation of Labor Income Processes," U. Texas Working Paper.

HEATHCOTE, J., K. STORESLETTEN, AND G. VIOLANTE (2006): "The Macroeconomic Implications of Rising Wage Inequality in the United States," NYU Working Paper.

JAIMOVICH, N. (2006): "Firm Dynamics, Markup Variations, and the Business Cycle," Stanford University Working Paper.

KRUEGER, D., AND F. PERRI (2005): "Public versus Private Risk Sharing," Working Paper.

KRUSELL, P., AND A. SMITH (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5).

LORENZONI, G. (2006): "A Theory of Demand Shocks," Working Paper.

LUCAS, R. (1987): *Models of Business Cycles.* Oxford: Basil Blackwell.

LUSTIG, H., AND S. VAN NIEUWERBURGH (2005): "Housing Collateral, Consumption Insurance and Risk Premia: An Empirical Perspective," *Journal of Finance*, 60(3), 1167–1219.

MURPHY, K., A. SHLEIFER, AND R. VISHNY (1989): "Building Blocks of Market-Clearing Models," *NBER Macroeconomics Annual*, pp. 247–286.

NEVO, A. (2001): "Measuring Market Power in the Ready-to-eat Cereal Industry," *Econometrica*, 69(2), 307–342.

PORTIER, F. (1995): "Business Formation and Cyclical Markups in the French Business Cycle," *Annales D'Economie et de Statistique*, 37-38, 411–440.

RAMPINI, A. (2004): "Entrepreneurial Activity, Risk, and the Business Cycle," *Journal of Monetary Economics*, 51, 555–573.

RIOS-RULL, J. (1996): "Life-Cycle Economies and Aggregate Fluctuations," *Review of Economic Studies*, 63, 465–490.

ROTEMBERG, J., AND M. WOODFORD (1999): "The Cyclical Behavior of Prices and Costs," in *Handbook of Marcoeconomics*, ed. by J. Taylor, and M. Woodford. Elsevier.

STOCK, J., AND M. WATSON (1999): "Business Cycle Fluctuations in US Macroeconomic Time Series," in *Handbook of Marcoeconomics*, ed. by J. Taylor, and M. Woodford. Elsevier.

STORESLETTEN, K., C. TELMER, AND A. YARON (2004): "Cyclical Dynamics in Idiosyncratic Labor Market Risk," *Journal of Political Economy*, 112(3), 695–717.