Macroeconomic Dynamics, **20**, 2016, 1101–1125. Printed in the United States of America.
doi:10.1017/S1365100514000753

PRICE-SETTING WITH UNOBSERVABLE ELASTICITIES OF DEMAND: THE BUSINESS-CYCLE EFFECTS OF HETEROGENEOUS EXPECTATIONS

CHRISTIAN JENSEN
University of South Carolina

In a dynamic stochastic general equilibrium model with monopolistic competition and flexible prices, we assume that producers must estimate their demand elasticities, which leads to heterogeneous expectations because of idiosyncratic shocks. I argue that these expectations shape firms' perceptions of relative prices, market shares, and individual demand elasticities, thereby distorting their price-setting and production. This model concludes that discarding the conventional assumption of known and exogenous demand elasticity generates business cycle fluctuations indistinguishable from those produced by traditional productivity shocks.

Keywords: Price-Setting, Heterogeneous Expectations, Misallocation, Productivity

1. INTRODUCTION

The mechanisms by which prices adjust to clear markets are not well understood. Even with monopolistic competition, where prices are set by sellers to maximize profits, our models tend to oversimplify, assuming that price-setters merely apply a known exogenous markup to their nominal production costs. In reality, price-setting is probably not this simple. For example, Bresnahan and Reiss (1991), Martins et al. (1996), and Campbell and Hopenhayn (2005) provide empirical evidence in support of markups reacting negatively to entry and the number of competitors. In addition, the optimal markup can depend on competitors' prices, and sellers might not have perfect information about this, or all other variables that affect optimal price-setting, when they decide their prices. The convenience of assuming exogenously given markups is that in the absence of nominal rigidities, relative prices are set optimally through the unsynchronized actions of individual producers, as a result of the coordination of their nominal prices through production costs.[1] However, moving away from this simplistic framework, even just by assuming that the exogenously optimal markups are not known by price-setters,

Address correspondence to: Christian Jensen, Department of Economics, University of South Carolina, 1014 Greene Street, Columbia, SC 29208, USA; e-mail: cjensen@alumni.cmu.edu.

 1365-1005/15

but must be estimated, can have a large enough effect on relative prices to have a significant impact on aggregate output. Moreover, I show that these distortions could explain a considerable part of the fluctuations observed in real GDP and be indistinguishable from the productivity shocks used to explain business-cycle fluctuations in the tradition of Kydland and Prescott (1982).

Idiosyncratic shocks to the production technology for final goods, or consumers' tastes, generate heterogeneity in a priori identical monopolistically competing intermediate-good producers' perceived elasticities of demand, distorting their relative prices. This, in turn, affects the cost-minimizing composition of final goods. By substituting intermediate goods that are priced too low for those that are priced too high, final-good producers keep costs down, sacrificing productivity. In a more general setup that captures the strategic interaction between competitors and the interdependence across markets, a monopolistic seller's optimal price would depend on the simultaneous pricing decisions of other sellers and the contemporary size of the market, which, in turn, depend on the concurrent actions of all sellers. This interdependence and the simultaneity of the price-setting imply that pricing decisions must be based on expectations of marketwide, or even economywide, aggregate variables. When such expectations differ across price-setters, relative prices of intermediate goods become distorted, affecting the productivity with which they are combined into final goods. As the dispersion in expectations of aggregate variables fluctuates over time with changing levels of uncertainty and shock heterogeneity, the productivity with which final goods are produced also varies, making aggregate output fluctuate.[2]

Surveys show that expectations of economic aggregates differ across both households and professional forecasters. What could generate such heterogeneity in a model with rational expectations?[3] In practice, dispersion arises from forecasts being based on different models, ideas, and idiosyncratic observations. There could be many reasons that such diverse practices, including nonrational ones, persist, but without doubt an important factor is that there is still too much we do not completely understand about how the aggregate economy works, in part because of poor, or nonexistent, measures of key variables. For example, using the neoclassical model to forecast aggregate output requires projecting total factor productivity, but without a good understanding of the underlying causes of productivity shocks, these have to be estimated from Solow residuals. This, in turn, requires having a reliable measure of the capital stock, which we lack. As a result, GDP forecasts tend to be based on ad hoc approaches that to a great extent ignore macroeconomic theory. In the model presented here, the capital stock is assumed to be perfectly observable, and can, together with the labor input, be used to deduce aggregate total factor productivity. However, its value depends on the efficiency with which each of the intermediate goods is produced and with which they are combined into final goods, and on the markups applied by each of their producers, which we assume is never measured across all the heterogeneous firms. Hence, although individual price-setters can easily obtain a history of aggregate productivity, we assume they cannot observe all the underlying variables affecting it, which would require obtaining a cross section of all their competitors. Consequently, they must base their expectations of aggregate productivity on its past values, which may or may not be rational, and on estimates of the underlying processes generated with their own private history of shocks. Because each price-setter will have a unique history of idiosyncratic shocks, they will produce different estimates of the processes generating these, even if they are a priori identical for all. Consequently, they will have heterogeneous expectations of aggregate productivity and of all other variables that depend on it.[4]

As in the misperceptions model of Lucas (1972), it is incomplete information that distorts relative prices and output in our setup. Although the problem is not caused by (random monetary) policy in our model, there is still a role for policy in improving the situation by contributing toward homogenizing expectations. Though factor prices serve as a coordination device for simultaneous price-setters, helping these achieve their optimal relative prices, this coordination is insufficient when the optimal markup depends on the simultaneous actions of one's competitors, and thus, on contemporary aggregate variables yet to be realized. As a result, there is scope for improving welfare by coordinating expectations about such variables, thereby minimizing the distortions to relative prices, and raising productivity. In particular, we argue that this is relevant to the aggregate price and output levels, because sellers cannot determine their elasticity of demand, and thus their optimal price, without knowing how their relative prices and market shares evolve.

In order for heterogeneous expectations to generate business cycles, the degree of disparity in expectations must vary over time. As heterogeneity increases, the dispersion in markups becomes larger, increasing distortions to relative prices, thus lowering aggregate total factor productivity. When heterogeneity falls, dispersion in markups falls, reducing distortions to relative prices and raising aggregate productivity. In terms of aggregate variables, Mankiw et al. (2003) find that the dispersion in inflation expectations vary greatly over time. With respect to the dispersion in markups, the lack of good measures at business-cycle frequencies is a problem. Survey data of plants, or even industries, are at best annual. In addition, markups are not measured directly, but must be approximated based on sales prices and estimated production costs. Domowitz et al. (1986a, Table 1) report that the standard deviation of price–cost margins, a commonly used estimate of markups, across 284 manufacturing industries dropped from 0.058 in 1958–1965 to 0.033 in 1974–1981. Domowitz et al. (1986b, Table I) show that from 1958 to 1981, the difference between the average markup of the 20% most concentrated industries in U.S. manufacturing and that of the 20% least concentrated varied from 0.064 to 0.152, increasing or decreasing by as much as 0.04 in a year. Studying the persistence in markups and prices across plants, Roberts and Supina (2000) argue that there is an important role for idiosyncratic shocks in generating markup variation. Eisfeldt and Rampini (2006) find that the cross-sectional productivity dispersion across firms behaves countercyclically. Kehrig (2011) confirms this finding, showing that the dispersion in total factor productivity levels across U.S. manufacturing plants is greater in recessions than in booms, which is evidence of the productivity mix varying over the business cycle (possibly because of increased markup dispersion). Other evidence of cyclical reallocation among plants is summarized by Gabaix (2011): "Most estimates of plant-level volatility find very large volatilities of sales and employment, with an order of magnitude $\sigma = 30 - 50\%$ per year [e.g., Davis et al. (1996); Caballero et al. (1997)]. Also, the volatility of firm size in Compustat is very large, 40% per year [Comin and Mulani (2006)]."

Our dynamic general equilibrium model builds on that of Blanchard and Kiyotaki (1987). It consists of an infinite number of a priori identical monopolistically competing producers that rent capital and labor from households in competitive factor markets to produce differentiated intermediate goods that households purchase to compose final goods. The production side of the economy is presented in the next section, including intermediate-good producers' price-setting decisions, which are initially functions of the expected values of their exogenous elasticities of demand, which they need to estimate. The following two sections describe the households, which face a standard dynamic decision problem, and the equilibrium conditions for the model economy, respectively. The subsequent section motivates how idiosyncratic shocks can make a priori identical intermediate-good producers obtain different estimates for their exogenous demand elasticities and the distribution of these. The following sections study the impact this heterogeneity has on aggregate output for different values of the elasticity of substitution between intermediate goods. This is first done for an inelastic labor supply (steady state), as this requires only specifying a distribution for the expected demand elasticities across intermediate-good producers and the weight of capital versus labor in the production of intermediate goods (α). The analysis is then repeated for the case where households adjust their labor supply optimally over time, which requires a more detailed specification and calibration of the model. Although assuming exogenously given elasticities of demand that price-setters must estimate yields a relatively simple and familiar framework for studying the effects of heterogeneous expectations, these elasticities are in reality likely to be endogenous and to depend on the contemporary shocks and actions of all competitors. The eighth section illustrates this, and argues that heterogeneity in perceived elasticities can arise from aggregate uncertainty, that is, from expectations of contemporary aggregates differing across price-setters. We conclude that variations in the dispersion of heterogeneous expectations produced by changes in the degree of uncertainty can contribute significantly to business cycle fluctuations and be almost indistinguishable from the productivity shocks traditionally used to explain these.

2. PRODUCERS

Each of the continuum of measure one of identical households produces y units of the final good by combining a continuum of differentiated intermediate goods

x_i indexed by $i \in [0, 1]$, so that

$$y_t = \left(\int_0^1 (\gamma_i x_i)^{\frac{\theta-1}{\theta}} \, di \right)^{\frac{\theta}{\theta-1}}, \tag{1}$$

where $\theta \in (1, \infty)$ is the elasticity of substitution between any two intermediate goods. This is the Dixit–Stiglitz aggregator modified to incorporate productivity shocks γ_i that can change the relative weight of each intermediate good in the production of the final good, as well as the general productivity of intermediate goods in the production of the final good; see Dixit and Stiglitz (1977).[5] Assuming that intermediate goods are the only inputs required to produce final goods, each household chooses the optimal mix of theses to minimize the cost of providing final goods by solving

$$\min_{\{x_i\}_{i=0}^{1}} \int_0^1 P_i x_i di, \tag{2}$$

subject to the production function (1), where P_i is the price of intermediate good i. The resulting demand for intermediate good i from each of the households is

$$x_i = \left(\frac{P_i}{P} \right)^{-\theta} \gamma_i^{\theta-1} y \tag{3}$$

for all $i \in [0, 1]$, where all variables are known by the household, because it is assumed to know its demand y for the final good, the prices P_i of the intermediate goods it buys, the shocks γ_i, and the marginal cost of producing the final good,

$$P = \left[\int_0^1 \left(\frac{P_i}{\gamma_i} \right)^{1-\theta} di \right]^{\frac{1}{1-\theta}}, \tag{4}$$

which is derived by inserting the demand for intermediate goods (3) into the production function for final goods (1). Because all households are identical, they compose identical final goods at identical cost, and because they can all produce the good, its market price equals its cost of production (4). Aggregating intermediate-good demands (3) across all households, we find the aggregate demand for intermediate good i to be

$$X_i = \left(\frac{P_i}{P} \right)^{-\theta} \gamma_i^{\theta-1} Y, \tag{5}$$

where Y is the aggregate demand for final goods.

Intermediate-good producer i finds the optimal mix of inputs, capital k_i, labor n_i, and land l_i to minimize its production costs by solving

$$\min_{k_i, n_i, l_i} Rk_i + Wn_i + Fl_i, \tag{6}$$

subject to its production technology

$$X_i = z_i k_i^{\alpha} n_i^{1-\alpha-\nu} l_i^{\nu}, \tag{7}$$

https://doi.org/10.1017/S1365100514000753 Published online by Cambridge University Press

where W is the nominal wage, R is the nominal rental rate of capital, F is the nominal rental rate of land, $\alpha \in (0, 1)$, $\nu \in (0, 1)$, and z_i is an exogenous productivity shock known by producer i, but no one else.[6]

The resulting first-order conditions yield its factor demands,

$$k_i = \alpha \frac{\lambda_i X_i}{R}, \tag{8}$$

$$n_i = (1 - \alpha - \nu) \frac{\lambda_i X_i}{W}, \tag{9}$$

$$l_i = \nu \frac{\lambda_i X_i}{F}, \tag{10}$$

where

$$\lambda_i = \frac{1}{z_i} \left(\frac{R}{\alpha}\right)^{\alpha} \left(\frac{W}{1 - \alpha - \nu}\right)^{1-\alpha-\nu} \left(\frac{F}{\nu}\right)^{\nu} \tag{11}$$

is the marginal cost of producing intermediate good i. In addition to choosing the cost-minimizing input mix, intermediate-good producer i needs to price its good. It does so by choosing the price P_i that maximizes its expected profits given the demand it faces (5), and thus solves

$$\max_{P_i} E_i \left[(P_i - \lambda_i) \left(\frac{P_i}{P}\right)^{-\theta} \gamma_i^{\theta-1} Y \right], \tag{12}$$

where E_i is the expectation operator, which is necessary because the intermediate-good producer i cannot observe the shock γ_i, the elasticity θ, aggregate demand Y, or the aggregate price level P. Exploiting the fact that it can observe the demand X_i for its good and the marginal cost λ_i of producing it, the profit-maximizing condition can be written as

$$X_i \left(1 - \frac{P_i - \lambda_i}{P_i} E_i \theta \right) = 0, \tag{13}$$

so intermediate-good producer i continuously adjusts its price P_i until this condition is satisfied, which makes the optimal price

$$P_i = \lambda_i \frac{E_i \theta}{E_i \theta - 1}, \tag{14}$$

a gross markup $\frac{E_i\theta}{E_i\theta - 1} \in (1, \infty)$ on the marginal cost of production λ_i.

Inserting the optimal pricing condition (14) into the price aggregator (4), after substituting for the marginal cost of production (11), yields the price level

$$P = \left(\frac{R}{\alpha}\right)^{\alpha} \left(\frac{W}{1 - \alpha - \nu}\right)^{1-\alpha-\nu} \left(\frac{F}{\nu}\right)^{\nu} \left[\int_0^1 \left(\frac{1}{\gamma_i z_i} \frac{E_i \theta}{E_i \theta - 1} \right)^{1-\theta} di \right]^{\frac{1}{1-\theta}}, \tag{15}$$

https://doi.org/10.1017/S1365100514000753 Published online by Cambridge University Press

and thus the relative price

$$\frac{P_i}{P} = \frac{\frac{1}{z_i}\frac{E_i\theta}{E_i\theta-1}}{\left[\int_0^1 \left(\frac{1}{\gamma_i z_i}\frac{E_i\theta}{E_i\theta-1}\right)^{1-\theta} di\right]^{1/(1-\theta)}}. \tag{16}$$

Substituting this relative price into the demand function for intermediate good i (5) and substituting the resulting equation and the marginal cost of production (11) into the factor demands (8), (9), and (10), and aggregating over all intermediate-good producers, we find the aggregate demands for capital, labor, and land,

$$K = \left(\frac{R}{\alpha}\right)^{\alpha-1}\left(\frac{W}{1-\alpha-\nu}\right)^{1-\alpha-\nu}\left(\frac{F}{\nu}\right)^{\nu} Y \frac{\int_0^1 \left(\frac{1}{\gamma_i z_i}\right)^{1-\theta}\left(\frac{E_i\theta}{E_i\theta-1}\right)^{-\theta} di}{\left[\int_0^1 \left(\frac{1}{\gamma_i z_i}\frac{E_i\theta}{E_i\theta-1}\right)^{1-\theta} di\right]^{-\theta/(1-\theta)}} \tag{17}$$

$$N = \left(\frac{R}{\alpha}\right)^{\alpha}\left(\frac{W}{1-\alpha-\nu}\right)^{-\alpha-\nu}\left(\frac{F}{\nu}\right)^{\nu} Y \frac{\int_0^1 \left(\frac{1}{\gamma_i z_i}\right)^{1-\theta}\left(\frac{E_i\theta}{E_i\theta-1}\right)^{-\theta} di}{\left[\int_0^1 \left(\frac{1}{\gamma_i z_i}\frac{E_i\theta}{E_i\theta-1}\right)^{1-\theta} di\right]^{-\theta/(1-\theta)}}, \tag{18}$$

$$L = \left(\frac{R}{\alpha}\right)^{\alpha}\left(\frac{W}{1-\alpha-\nu}\right)^{1-\alpha-\nu}\left(\frac{F}{\nu}\right)^{\nu-1} Y \frac{\int_0^1 \left(\frac{1}{\gamma_i z_i}\right)^{1-\theta}\left(\frac{E_i\theta}{E_i\theta-1}\right)^{-\theta} di}{\left[\int_0^1 \left(\frac{1}{\gamma_i z_i}\frac{E_i\theta}{E_i\theta-1}\right)^{1-\theta} di\right]^{-\theta/(1-\theta)}}, \tag{19}$$

respectively. Because producers' decisions are not dynamic, the time subscripts have been suppressed throughout this section.

3. CONSUMERS

In addition to effortlessly creating final goods, households rent their labor N, capital K, and land L to the collectively owned intermediate-good producers in order to provide for consumption C and the accumulation of assets: physical capital K, money M. and bonds B. Because households are assumed to be identical, aggregation is trivial, so we focus on aggregates directly. Putting money into the utility function u as a short cut, each of the continuum of measure one of identical households solves the dynamic problem

$$\max_{C_t, N_t, K_t, B_t, M_t} E_0 \sum_{t=0}^{\infty} \beta^t u\left(C_t, 1-N_t, \frac{M_t}{P_t}\right), \tag{20}$$

subject to

$$K_t + \frac{B_t}{P_t} + \frac{M_t}{P_t} + C_t = \frac{W_t}{P_t} N_t + \frac{R_t}{P_t} K_{t-1} + \frac{F_t}{P_t} + (1-\delta) K_{t-1} + \frac{(1+\Re_{t-1}) B_{t-1}}{P_t} + \frac{M_{t-1}}{P_t} + \frac{\Pi_t}{P_t} + S_t, \tag{21}$$

given a discount rate $\beta \in (0, 1)$, a depreciation rate $\delta \in (0, 1)$, and initial conditions K_{t-1}, B_{t-1}, M_{t-1}, and $\Re_{t-1}$, where W_t/P_t is the real wage, R_t/P_t is the real rental rate of capital, $\frac{F_t}{P_t}$ is the real rental cost of land, $\Re_t$ is the nominal interest rate on bonds, Π_t is profits from the production of intermediate goods, S_t is transfers from the government, and P_t is the aggregate price level, equal to the price of the final good. To simplify, the supply of land is normalized to one. The first-order conditions are given by the budget constraint (21) and

$$u_2'(C_t, 1-N_t, m_t) = u_1'(C_t, 1-N_t, m_t) \frac{W_t}{P_t}, \tag{22}$$

$$u_1'(C_t, 1-N_t, m_t) = \beta E_t u_1'(C_{t+1}, 1-N_{t+1}, m_{t+1}) \left(\frac{R_{t+1}}{P_{t+1}} + 1 - \delta \right), \tag{23}$$

$$u_1'(C_t, 1-N_t, m_t) = (1+\Re_t) \beta E_t u_1'(C_{t+1}, 1-N_{t+1}, m_{t+1}) \left(\frac{P_{t+1}}{P_t} \right)^{-1}, \tag{24}$$

$$u_1'(C_t, 1-N_t, m_t) = u_3'(C_t, 1-N_t, m_t) + \beta E_t u_1'(C_{t+1}, 1-N_{t+1}, m_{t+1}) \left(\frac{P_{t+1}}{P_t} \right)^{-1}, \tag{25}$$

where $m_t = M_t/P_t$ is the demand for real money balances.

4. EQUILIBRIUM

Exploiting the government budget constraint

$$P_t S_t + P_t G_t = M_t - M_{t-1} + B_t - (1+\Re_{t-1}) B_{t-1}, \tag{26}$$

and that total profits for intermediate-good producers are

$$\Pi_t = \int_0^1 P_{i,t} z_{i,t} k_{i,\,t-1}^{\alpha} n_{i,t}^{1-\alpha-\nu} l_{i,t}^{\nu} di - R_t K_{t-1} - W_t N_t - F_t, \tag{27}$$

households' budget constraints (21) simplify to

$$K_t + C_t + G_t = Y_t + (1-\delta) K_{t-1}, \tag{28}$$

where

$$Y_t = \int_0^1 \frac{P_{i,t}}{P_t} z_{i,t} k_{i,t-1}^{\alpha} n_{i,t}^{1-\alpha-\nu} l_{i,t}^{\nu} di \tag{29}$$

is the real value of production.[7] Setting aggregate demand for land (19) equal to the inelastic unitary supply yields aggregate output

$$Y_t = K_{t-1}^{\alpha} N_t^{1-\alpha} \frac{\left[\int_0^1 \left(\gamma_{i,t} z_{i,t} \frac{E_{i,t}\theta - 1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{\theta/(\theta-1)}}{\int_0^1 \left(\gamma_{i,t} z_{i,t}\right)^{\theta-1} \left(\frac{E_{i,t}\theta-1}{E_{i,t}\theta}\right)^{\theta} di} \tag{30}$$

after exploiting that the aggregate demands for factors of production (17)–(19) imply that

$$\frac{R_t}{F_t} = \frac{\alpha}{\nu K_{t-1}}, \tag{31}$$

$$\frac{W_t}{F_t} = \frac{1-\alpha-\nu}{\nu N_t}, \tag{32}$$

which guarantee an optimal input mix in the production of intermediate goods. Combining these two conditions with the one for the price level (15) yields

$$\frac{R_t}{P_t} = \alpha K_{t-1}^{\alpha-1} N_t^{1-\alpha-\nu} \left[\int_0^1 \left(\gamma_{i,t} z_{i,t} \frac{E_{i,t}\theta - 1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{1/(\theta-1)}, \tag{33}$$

$$\frac{W_t}{P_t} = (1-\alpha-\nu) K_{t-1}^{\alpha} N_t^{-\alpha-\nu} \left[\int_0^1 \left(\gamma_{i,t} z_{i,t} \frac{E_{i,t}\theta - 1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{1/(\theta-1)}, \tag{34}$$

$$\frac{F_t}{P_t} = \nu K_{t-1}^{\alpha} N_t^{1-\alpha-\nu} \left[\int_0^1 \left(\gamma_{i,t} z_{i,t} \frac{E_{i,t}\theta - 1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{1/(\theta-1)}, \tag{35}$$

which are the real rental rates and the real wage. Finally, the equilibrium price level is determined by the demand and supply for money,

$$P_t = \frac{M_t}{m_t}. \tag{36}$$

Because of the lack of significance of land as a source of fluctuations, we let ν converge toward zero, so that land is eliminated from the model henceforth. Heterogeneous expectations among intermediate-good producers about their demand elasticities $E_{i,t}\theta$ directly affect aggregate output (30) and the real payments to the factors of production, capital (33) and labor (34), and all other aggregate variables indirectly through these.

https://doi.org/10.1017/S1365100514000753 Published online by Cambridge University Press

TABLE 1. Effect of σ_γ on different percentile of markup relative to the median

σ_γ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55
5th	.92	.86	.80	.75	.71	.67	.63	.60	.57	.55	.53
10th	.94	.89	.84	.80	.76	.72	.69	.66	.63	.61	.59
25th	.97	.94	.91	.88	.86	.83	.81	.79	.77	.75	.73
75th	1.03	1.07	1.11	1.16	1.20	1.25	1.31	1.37	1.44	1.51	1.59
90th	1.07	1.15	1.24	1.34	1.47	1.62	1.81	2.1	2.4	2.8	3.4
95th	1.09	1.20	1.33	1.49	1.70	1.97	2.4	2.9	3.9	5.6	10.5

5. DISPERSION

As the preceding equilibrium conditions show, there is no way to deduce the value of the elasticity θ from aggregate data without knowing the distributions of $\gamma_{i,t}$, $z_{i,t}$, and $E_{i,t}\theta$ across all producers i. Because these distributions can change over time, in particular the one for expectations, $E_{i,t}\theta$, they are not likely to be known. However, when θ is constant over time, intermediate-good producer i can estimate it from the demand equation (5), using past data on aggregate output Y and the price level P, combined with past realizations of its own price P_i and sales X_i. Applying the logarithm to both sides of its demand function (5) yields the linear equation

$$\ln \frac{X_i}{Y} = -\theta \ln\left(\frac{P_i}{P}\right) + (\theta - 1)\ln \gamma_i, \tag{37}$$

where $(\theta - 1)\ln \gamma_i$ is an unobservable error term, because the γ_i-shocks are never observed. Whatever the standard deviation of $\ln \gamma_i$ is, the shock to the demand equation will be magnified by a factor of $\theta - 1$. Hence, the higher the elasticity θ, the more imprecise its estimate will be, by a factor of $(\theta - 1)$. Because the γ_i -shocks are heterogeneous, the estimates of θ will differ across firms. Borrowing from this, we assume in our following simulations that intermediate-good producers' expectational errors $E_i\theta - \theta$ are independently and identically distributed across producers according to a normal distribution with mean zero and standard deviation $(\theta - 1)\,\sigma_\gamma$, where σ_γ is the standard deviation of $\ln \gamma_i$ across firms. Having the standard error of the forecasts increase with θ is necessary to generate dispersion in intermediate-good producers' markups also for larger values of θ. In addition, it makes the distribution of the markups $1/(E_i\theta - 1)$ relative to the median markup $1/(\theta - 1)$ be independent of θ. This is illustrated in Table 1, which shows how the 5th, 10th, 25th, 75th, 90th, and 95th percentile markups are distributed around the median for different values of σ_γ, reporting these as a fraction of the median markup.

The productivity shocks γ_i and z_i are assumed to be independently distributed across firms according to a normal distribution with a unit mean. To simplify, these shocks are also assumed to be independent of the expectational errors $E_i\theta - \theta$.[8]

6. INELASTIC LABOR SUPPLY (STEADY STATE)

Although discerning all the effects of dispersion requires solving the dynamic model, the implications for equilibrium real output and factor payments can be obtained without doing so when the labor supply is perfectly inelastic. Hence, the present section ignores households' utility-maximizing condition for labor supply (22) and instead assumes that N_t equals some constant value N. The steady state with constant dispersion is a special case of this, so the results in this section also represent the long-run effects when the labor supply is elastic.

A mean-preserving spread of productivity, $z_{i,t}\gamma_{i,t}$, across firms would have a positive effect on both equilibrium output and the real return on capital, as a result of inputs flowing from low-productivity to high-productivity producers. This effect is greater the higher the elasticity θ, that is, the easier it is to substitute between intermediate goods. Because our primary interest is in determining the effects of heterogeneous expectations (of perceived elasticities $E_{i,t}\theta$), we control for the effects of dispersion in productivity, while at the same time allowing these two types of heterogeneity to interact. To do so, our benchmark assumes that the perceived elasticities are constant across firms, $E_{i,t}\theta = \theta$ for all i, while maintaining the heterogeneity in terms of the productivity shocks $z_{i,t}$ and $\gamma_{i,t}$.

Dividing output with dispersion in perceived elasticities (30) by that without ($E_{i,t}\theta = \theta$) yields the fraction

$$\bar{r}_t = \frac{\dfrac{\left[\int_0^1 \left(\gamma_{i,t} z_{i,t} \frac{E_{i,t}\theta - 1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{\theta/(\theta-1)}}{\int_0^1 (\gamma_{i,t} z_{i,t})^{\theta-1}\left(\frac{E_{i,t}\theta-1}{E_{i,t}\theta}\right)^{\theta} di}}{\left[\int_0^1 (\gamma_{i,t} z_{i,t})^{\theta-1} di\right]^{1/(\theta-1)}} \leq 1 \tag{38}$$

by which total factor productivity (TFP) and output are reduced because of such heterogeneity for any given levels of capital and employment. The effect of dispersion in perceived elasticities is to immediately reduce aggregate output, independent of whether or not the perceived elasticities are centered around the true value; only the dispersion matters. The intuition is that because producers face the same elasticity θ, disparities in their beliefs $E_{i,t}\theta$ make them choose suboptimal relative prices, which in turn make the composition of the final good inefficient, resulting in less of it being produced. Errors in beliefs that are homogeneous across producers have no impact (given the inelastic supply of the factors of production), because such errors do not affect relative prices [Lerner (1934)]. The greater the disparities in the estimated elasticities, $E_{i,t}\theta$, the greater the distortions to relative prices, and the less final good is produced. This negative impact on output is larger the smaller the value of $\theta \in (1, \infty)$, because this reduces the substitutability between intermediate goods, making it more difficult to adjust the mix of these in reaction to distortions to relative prices.

In addition to the immediate effect on output through the mix of intermediate goods, dispersion also affects output over time through capital accumulation. It does so not only through the wealth effect from the immediate drop in output, a rise in heterogeneity across intermediate-good producers' perceived elasticities that is expected to endure also affects capital accumulation through its rental rate. Comparing, as before, the rental rate with dispersion in expectations (33) with the one where $E_{i,t}\theta = \theta$,

$$\frac{R_t}{P_t} = \frac{\theta - 1}{\theta} \alpha K_{t-1}^{\alpha-1} \left[\int_0^1 \left(\gamma_{i,t} z_{i,t} \right)^{\theta-1} di \right]^{1/(\theta-1)}, \tag{39}$$

yields the multiple

$$\bar{q}_t = \frac{\left[\int_0^1 \left(\gamma_{i,t} z_{i,t} \frac{E_{i,t}\theta - 1}{E_{i,t}\theta} \right)^{\theta-1} di \right]^{1/(\theta-1)}}{\frac{\theta-1}{\theta} \left[\int_0^1 \left(\gamma_{i,t} z_{i,t} \right)^{\theta-1} di \right]^{1/(\theta-1)}}, \tag{40}$$

which measures the immediate impact dispersion in perceived elasticities $E_{i,t}\theta$ has on the equilibrium real rental rate of capital (and the real wage) for a given capital level. It can be greater or smaller than one, because the effect of dispersion can be positive or negative. From the preceding, we know that there is a negative effect from the inefficient mix of intermediate goods that dispersion in relative prices gives rise to, which makes capital be used less efficiently in producing the final good, thus reducing its rental rate. In addition, dispersion makes some firms apply a markup that is higher, and others apply a markup that is lower, than they otherwise would. This affects the rental rate, because the higher a markup an intermediate-good producer applies, the less capital it employs, as the equation for factor demand (8) shows, thus reducing the demand for capital and its rental rate. Whether this effect though the markups is positive or negative depends on the skewness of the distribution of the markups, and because firms that apply low markups become larger than those that apply high ones, a negative effect requires a positive skewness. In our case, this positive skewness arises due to the nonlinearity of the markups in terms of the perceived elasticities, and because the perceived elasticities $E_{i,t}\theta$ are required to be greater than one for the corresponding markups to be positive.[9] Even errors in beliefs that are homogeneous across firms have an impact on the rental rate, as they affect the average markup, and thus the demand for capital.

Although the immediate effect that dispersion in beliefs $E_{i,t}\theta$ has on aggregate output with an inelastic labor supply can be measured through the ratio $\bar{r}_t$, the long-run one, assuming the change in dispersion is permanent, needs to take into account the impact on the capital stock induced by the change in its rental rate. In a steady state with constant productivity and beliefs, so that $E_{i,t}\theta = E_i\theta$ and $\gamma_{i,t} z_{i,t} = \gamma_i z_i$, implying that the dispersion across intermediate-good producers is constant over time, and where, in addition, $\mu_t = \mu$, $G_t = G$, and $N_t = N$,

https://doi.org/10.1017/S1365100514000753 Published online by Cambridge University Press

aggregate output is

$$Y = N \left(\frac{\frac{1}{\beta} - 1 + \delta}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} \frac{\left[\int_0^1 \left(\gamma_i z_i \frac{E_i\theta - 1}{E_i\theta} \right)^{\theta-1} di \right]^{\frac{\theta+(1-\theta)\alpha}{(1-\alpha)(\theta-1)}}}{\int_0^1 (\gamma_i z_i)^{\theta-1} \left(\frac{E_i\theta-1}{E_i\theta} \right)^{\theta} di}. \quad (41)$$

In an identical steady state without dispersion in perceived elasticities, so that $E_i\theta = \theta$, we have

$$Y = N \left(\frac{\frac{1}{\beta} - 1 + \delta}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} \left(\frac{\theta - 1}{\theta} \right)^{\frac{\alpha}{1-\alpha}} \left[\int_0^1 (\gamma_i z_i)^{\theta-1} \, di \right]^{\frac{1}{(1-\alpha)(\theta-1)}}. \quad (42)$$

Thus, steady-state output with dispersion in beliefs $E_{i,t}\theta$ is a fraction

$$\bar{r} = \frac{\frac{\left[\int_0^1 \left(\gamma_i z_i \frac{E_i\theta-1}{E_i\theta} \right)^{\theta-1} di \right]^{\frac{\theta+(1-\theta)\alpha}{(1-\alpha)(\theta-1)}}}{\int_0^1 (\gamma_i z_i)^{\theta-1} \left(\frac{E_i\theta-1}{E_i\theta} \right)^{\theta} di}}{\left(\frac{\theta-1}{\theta} \right)^{\frac{\alpha}{1-\alpha}} \left[\int_0^1 (\gamma_i z_i)^{\theta-1} \, di \right]^{\frac{1}{(1-\alpha)(\theta-1)}}} \leq 1 \quad (43)$$

of what it would be in a steady state without such dispersion.

With a constant labor supply, we can quantify the impact dispersion in perceived elasticities $E_i\theta$ has on output through $\bar{r}_t$ immediately and through $\bar{r}$ in the long run by computing these. Doing so only requires specifying the distributions for productivity $\gamma_i z_i$ and perceived elasticities $E_i\theta$ and the true value of the elasticity θ, and, for the long-run effects, the value of α. Quantifying the impact of dispersion when the labor supply is elastic requires specifying the model in greater detail and solving out for the dynamics (see the next section).

A steady-state income share of 0.6 for labor implies that

$$\alpha = 1 - 0.6 \frac{\int_0^1 \left(\gamma_i z_i \frac{E_i\theta-1}{E_i\theta} \right)^{\theta-1} di}{\int_0^1 (\gamma_i z_i)^{\theta-1} \left(\frac{E_i\theta-1}{E_i\theta} \right)^{\theta} di}, \quad (44)$$

which in a steady state without dispersion means that

$$\alpha = 1 - 0.6 \frac{\theta}{\theta - 1}. \quad (45)$$

Ignoring the effects of dispersion, a strictly positive value of α requires $\theta > 2.5$.[10] Lowering the elasticity θ raises intermediate-good producers' markups, and thus the share of income that goes to profits. Because the income share of labor is fixed at 0.6, this means that the share of capital gets squeezed as more and more of the

https://doi.org/10.1017/S1365100514000753 Published online by Cambridge University Press

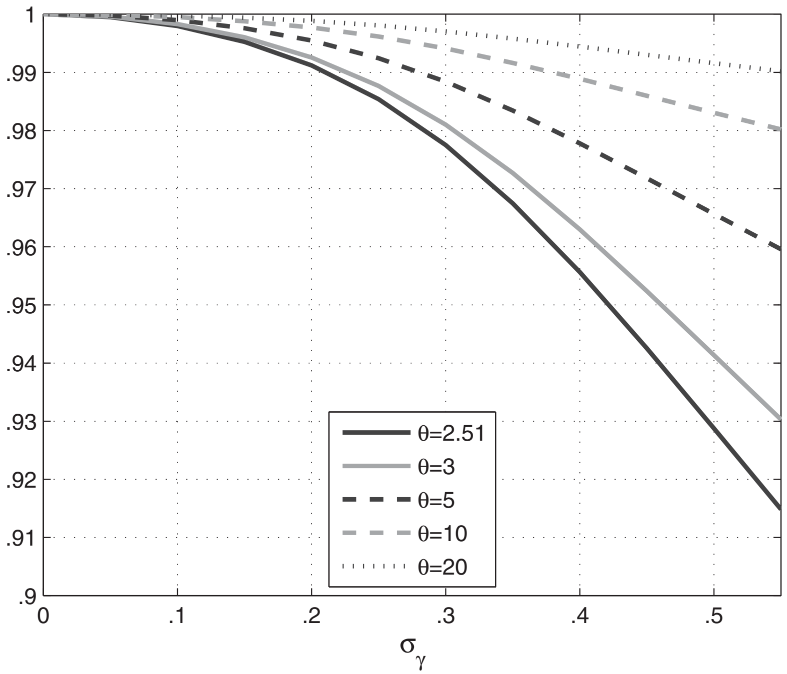

FIGURE 1. Instant effect of dispersion in perceived elasticities on TFP and output, $\bar{r}_t$ and r_t.

income goes to profits when the elasticity θ is lowered, making the importance of capital in production (α) fall.

Figure 1 plots $\bar{r}_t$, our measure of the immediate effect that dispersion in perceived elasticities $E_{i,t}\theta$ has on aggregate output, for given levels of capital and labor, as a function of the standard deviation σ_γ, defined earlier, for different values of θ. As expected, it shows that this effect is larger the lower the elasticity θ and the higher the dispersion σ_γ. At most, dispersion in perceived elasticities reduces output by about 8.5%, which is attainable only with minimal competition, $\theta = 2.51$. With a moderate level of competition, $\theta = 5$, the negative impact on output is at most 4.1%, and it is at most 1% with high competition, $\theta = 20$. Figure 2 plots $\bar{q}_t$, the measure of the immediate impact that dispersion in perceived elasticities has on the rental rate of capital, for given levels of capital and labor. With the distributions assumed earlier, dispersion in perceived elasticities reduces the rental rate; however, it does so by at most 5.5% for $\theta = 2.51$ and 0.5% for $\theta = 20$. Figure 3 plots $\bar{r}$, the measure of the impact that dispersion in perceived elasticities has on steady-state output. It shows that the impact on output is amplified by allowing capital to adjust, making the maximum reduction in output 8.7% for $\theta = 2.51$, 4.8% for $\theta = 5$, and 1.3% for $\theta = 20$. The effect

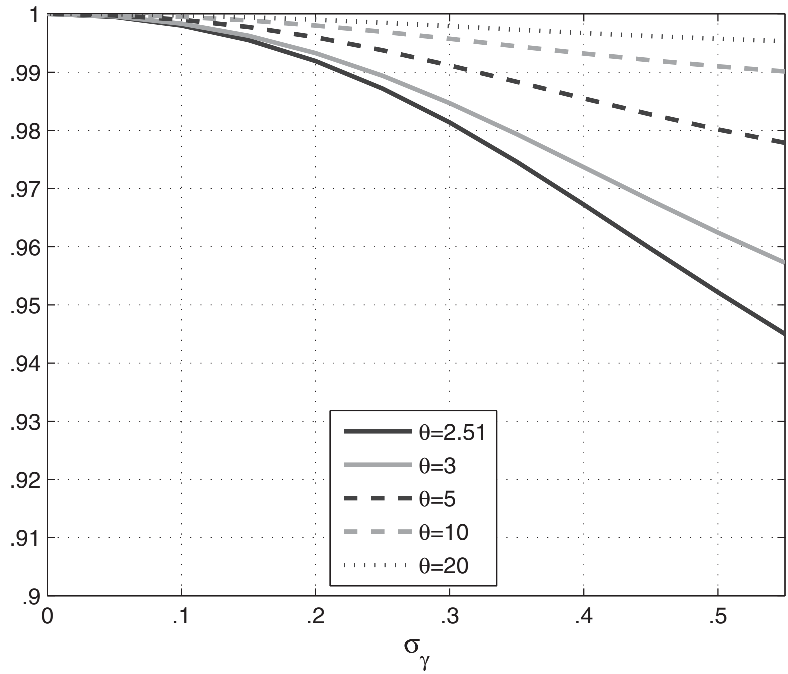

FIGURE 2. Instant effect of dispersion in perceived elasticity on factor prices, $\bar{q}_t$ and $\theta/(\theta-1)q_t$.

on output from letting capital adjust is greater for larger θ, even though this makes dispersion in the perceived elasticities have a smaller impact on the rental rate, because α is larger the higher θ is, making capital carry a larger weight in production.

Intuitively, the effects of dispersion should be greater when the labor supply is elastic. The reason is that dispersion in perceived elasticities has the same impact on the real wage as on the rental rate of capital, and therefore diminishes the labor supply, which adjusts immediately, instead of gradually, as capital does. This is quantified in the following section.

7. ELASTIC LABOR SUPPLY

To simplify the dynamic analysis, and because our main focus is on the impact of dispersion in perceived elasticities $E_{i,t}\theta$, we henceforth eliminate the dispersion in productivity by letting $\gamma_{i,t}z_{i,t} = \gamma_t z_t = Z_t$ for all i, noting that these two variables have identical effects on the aggregate variables of interest. This shuts down any interaction there might be between the two types of dispersion, but this interaction appears to be insignificant with a large enough number of firms.[11] As a result,

https://doi.org/10.1017/S1365100514000753 Published online by Cambridge University Press

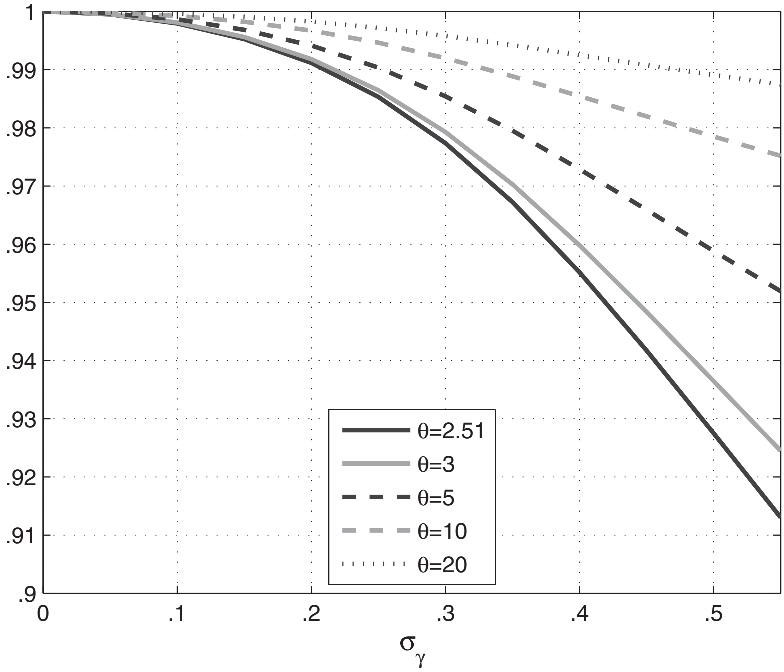

FIGURE 3. Effect of dispersion in perceived elasticities on steady-state output, $\bar{r}$.

aggregate output (30) simplifies to

$$Y_t = Z_t K_{t-1}^{\alpha} N_t^{1-\alpha} \frac{\left[\int_0^1 \left(\frac{E_{i,t}\theta-1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{\theta/(\theta-1)}}{\int_0^1 \left(\frac{E_{i,t}\theta-1}{E_{i,t}\theta}\right)^{\theta} di}, \quad \textbf{(46)}$$

whereas the real rental rate of capital and the real wage are

$$\frac{R_t}{P_t} = \alpha Z_t K_{t-1}^{\alpha-1} N_t^{1-\alpha} \left[\int_0^1 \left(\frac{E_{i,t}\theta - 1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{1/(\theta-1)}, \quad \textbf{(47)}$$

$$\frac{W_t}{P_t} = (1-\alpha)\, Z_t K_{t-1}^{\alpha} N_t^{-\alpha} \left[\int_0^1 \left(\frac{E_{i,t}\theta - 1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{1/(\theta-1)}, \quad \textbf{(48)}$$

https://doi.org/10.1017/S1365100514000753 Published online by Cambridge University Press

respectively. These are the only equilibrium conditions where dispersion enters directly, so we focus on these. In particular, we are interested in the fraction

$$r_t = \frac{\left[\int_0^1 \left(\frac{E_{i,t}\theta-1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{\theta/(\theta-1)}}{\int_0^1 \left(\frac{E_{i,t}\theta-1}{E_{i,t}\theta}\right)^{\theta} di} \leq 1 \tag{49}$$

and the factor

$$q_t = \left[\int_0^1 \left(\frac{E_{i,t}\theta-1}{E_{i,t}\theta}\right)^{\theta-1} di\right]^{1/(\theta-1)}, \tag{50}$$

where r_t measures the immediate impact of heterogeneous expectations on aggregate total factor productivity and output for given levels of capital K_{t-1}, work effort N_t, and productivity Z_t, whereas q_t measures the instantaneous impact of dispersion on the factor prices under the same conditions. Because of the negligible effects of the interaction between dispersion in productivity and in the perceived elasticities, $r_t \approx \bar{r}_t$ and $q_t \approx [(\theta-1)/\theta]\bar{q}_t$, so Figure 1 also represents r_t, whereas Figure 2 also applies for $[\theta/(\theta-1)]q_t$.

The approximation

$$q_t \simeq \left(1-\frac{1}{\theta}\right) r_t \tag{51}$$

is accurate for $\sigma_\gamma \leq 0.55$.[12] With it, we can write the equations for aggregate output (46), real rental rate (39), and real wage (48) as

$$Y_t = Z_t K_{t-1}^{\alpha} N_t^{1-\alpha} r_t, \tag{52}$$

$$\frac{R_t}{P_t} = \alpha Z_t K_{t-1}^{\alpha-1} N_t^{1-\alpha}\left(1-\frac{1}{\theta}\right) r_t, \tag{53}$$

and

$$\frac{W_t}{P_t} = (1-\alpha)\, Z_t K_{t-1}^{\alpha} N_t^{-\alpha}\left(1-\frac{1}{\theta}\right) r_t, \tag{54}$$

respectively. Because the productivity shock Z_t and the dispersion shock r_t enter the model in exactly the same way, their effects on the aggregate variables will be identical if they posses the same statistical properties.[13] The standard deviation of the productivity shock Z_t is typically assumed to be about 0.0225, implying that it is within $\pm 4.5\%$ of its steady-state value 95% of the time. Looking at Figure 1, it is evident that swings in r_t of this magnitude cannot be generated unless σ_γ is increased beyond 0.55. Hence, it would take an implausible amount of dispersion for such r_t-shocks to be the sole source of productivity shocks. Exactly what fraction of these could be dispersion shocks depends on θ, and how σ_γ changes over time, which determines the variance and autocorrelation of r_t.

With the utility function

$$u(C_t, 1 - N_t) = \frac{b\left[C_t^a (1 - N_t)^{1-a}\right]^{1-\phi} + (1 - b)\, m_t^{1-\phi}}{1 - \phi}, \tag{55}$$

where a measures the weight of consumption relative to leisure, $1 - b$ measures the weight of real money balances relative to consumption and leisure, and ϕ is the inverse of the intertemporal elasticity of substitution, we calibrate the model for $\beta = 0.989$, $\delta = 0.028$, $\phi = 0.5$, $G/Y = 0.2$, and $N = 0.3$; see Cooley and Hansen (1995) and Hansen (1997).[14] In addition, we choose α to yield an income share of 0.6 in a steady state (44) with constant dispersion σ_γ (discussed earlier). Log-linearizing around such a steady state, we can measure the impact that dispersion shocks r_t have on aggregate output, taking into account the reaction of the supply of labor and capital to changes in factor prices spurred by variations in the dispersion of perceived elasticities.

When dispersion shocks r_t are highly correlated over time, with an autocorrelation coefficient $\rho_r = 0.95$, we find that a 1% deviation in r_t from steady state makes contemporaneous aggregate output deviate about 1.44% from steady state for $\theta \in [5, \infty)$.[15] For lower values of the elasticity θ, the amplification of dispersion shocks r_t is somewhat lower: 1.41 for $\theta = 4$, 1.31 for $\theta = 3$, and 0.9 for $\theta = 2.51$. The reason is that though the effect of an r_t shock on aggregate output through productivity is the same for all values of θ, the impact through the inputs is lower for smaller θ, because this makes intermediate-good producers limit the quantity produced more tightly. These results are quite robust, in that for $\theta \in [3, \infty)$, achieving an output response below 1% or above 2% would require an extreme calibration for $\beta, \delta, \phi, G/Y, N$, and the income share of labor. The results are, however, fairly sensitive to the autocorrelation ρ_r of the shocks (discussed later). Hence, we find that when $\rho_r = 0.95$, taking into account the effect of an elastic labor supply at most allows doubling the effects seen in Figure 1, even though magnifying these by a factor of 0.9-1.44, depending on the value of θ, is more realistic. This implies that when $\sigma_\gamma \leq 0.55$, dispersion in the perceived elasticities $E_{i,t}\theta$ can reduce aggregate output by about 7.7% when $\theta = 2.51$, 9.2% when $\theta = 3$, 7.2% when $\theta = 4$, 5.8% when $\theta = 5$, 2.9% when $\theta = 10$, and 1.4% when $\theta = 20$. Assuming that the steady-state value of dispersion σ_γ in each case is such that the impact on output is half of the maximum impact, we find that fluctuations in dispersion, that is, changes in σ_γ between 0 and 0.55, can generate output deviations from steady state of $\pm 3.8\%$ when $\theta = 2.51$, $\pm 4.6\%$ when $\theta = 3$, $\pm 3.6\%$ when $\theta = 4$, $\pm 2.9\%$ when $\theta = 5$, $\pm 1.4\%$ when $\theta = 10$, and $\pm .7\%$ when $\theta = 20$. For comparison, Table 2 presents the cumulative deviations from trend of U.S. real GDP. It shows that from the first quarter of 1960 to the fourth quarter of 2011, real GDP was within $\pm 1.5\%$ of its trend 70.6% of the quarters, and within $\pm 3\%$ of its trend 94.1% of the time.

The autocorrelation coefficient ρ_r has a large impact on the effect r_t shocks have on output, because it determines their persistence, which shapes the labor

Table 2. U.S. real GDP deviations from Hodrick–Prescott (1997) trend, 1960.1–2011.4

% dev. trend	±.5	±1.0	±1.5	±2.0	±2.5	±3.0	±3.5	±4.0	±4.7
% of quarters	26.5	49.5	70.6	82.8	87.3	94.1	96.6	98.5	100

response. The more persistent the r_t shocks are, the less intertemporal substitution of labor they generate, reducing their impact on output. The calibration $\rho_r = 0.95$ is the standard one for productivity shocks, but it is unclear how persistent the dispersion shocks we have in mind should be. When $\rho_r = 0.5$, a 1% deviation in r_t from steady state makes aggregate output deviate about 1.12% from steady state with $\theta = 2.51$, 2.26% with $\theta = 3$, 2.15% with $\theta = 4$, 2.06% with $\theta = 5$, 1.89% with $\theta = 10$, and 1.82% with $\theta = 20$. In this case, dispersion in the perceived elasticities $E_{i,t}\theta$ can reduce aggregate output by at most 9.5% when $\theta = 2.51$, 15.8% when $\theta = 3$, 11.0% when $\theta = 4$, 8.3% when $\theta = 5$, 3.7% when $\theta = 10$, and 1.8% when $\theta = 20$. Assuming, as before, that the steady state is at the midpoint, fluctuations in dispersion, σ_γ, varying between 0 and 0.55 can generate output deviations from steady state of ±4.8% when $\theta = 2.51$, ±7.9% when $\theta = 3$, ±5.5% when $\theta = 4$, ±4.2% when $\theta = 5$, ±1.9% when $\theta = 10$, and ±0.9% when $\theta = 20$.

Despite the robustness mentioned earlier, the calibration affects the results in the following ways. Increasing the labor share amplifies the effects that dispersion shocks have on output, because it reduces the importance of capital, which cannot immediately react to shocks, and increases that of labor, which does adjust instantly. Increasing the discount rate β, or decreasing the depreciation rate δ, raises the steady-state capital stock, thus raising the steady-state marginal product of labor, which increases the impact that changes in the labor input have on output, thereby increasing the amplification of dispersion shocks. Similarly, lowering the steady-state value of the labor input N raises the steady-state marginal product of labor, and therefore also raises the amplification. Lowering ϕ raises the intertemporal substitution, making households vary their labor supply more in response to changes in the real wage, which is determined by the labor productivity and the dispersion shock. Because government spending G is exogenous and constant in the model, dispersion shocks have smaller effects the larger the government sector is.[16]

Although dispersion has a greater impact on aggregate output for lower values of the elasticity θ, we should keep in mind that in terms of the markups, the dispersion is also larger for these. It follows that obtaining comparable numbers for different values of θ might arguably require letting σ_γ reach higher values the larger θ is. This would make our results more similar across θ values. There would be no effect on the amplification of the dispersion shocks r_t. The magnitude of the fluctuations in aggregate output that dispersion shocks can generate would still depend crucially on the width of the range we allow for the dispersion σ_γ, the autocorrelation ρ_r, and the value of the elasticity θ.[17]

8. DEMAND ELASTICITIES IN GENERAL

Although the elasticity θ is exogenous, constant, and identical for all producers of intermediate goods, it is not possible to deduce, or estimate, its value from aggregate data without knowing the distribution of its perceived value $E_{i,t}\theta$ across all i. As a result, intermediate-good producers are forced to form their expectations of its value from estimates of the demand (37) for their good i. Due to unobservable heterogeneous shocks, the estimates they produce differ, which leads to dispersion in perceived elasticities. This is crucial, because the preceding results rely on this dispersion to distort relative prices of intermediate goods, which in turn affects aggregate productivity. With identical θ values for all producers, they could get data from each other to produce more accurate estimates. However, in a world where elasticities vary across producers, such information sharing would not be useful, and each would be stuck with its own estimation problems and estimates.

In a more general setup, the elasticity of demand faced by each producer would be endogenous, and depend on the idiosyncratic shocks and price-setting decisions of all its competitors, as well as aggregate demand Y. To see this, imagine that production of final good is given by

$$y = f\left(\{x_i\}_{i=0}^{1}, \{\gamma_i\}_{i=0}^{1}\right). \tag{56}$$

Minimizing the cost of producing the final good (2) now yields an aggregate demand function for intermediate good i,

$$X_i = D_i\left(\left\{\frac{P_i}{P}\right\}_{i=0}^{1}, \{\gamma_i\}_{i=0}^{1}, Y\right), \tag{57}$$

where P is the price of the final good, equal to its marginal cost of production. Maintaining the Cobb–Douglas production technology (7) for intermediate goods, the marginal cost of producing good i remains as before (11), and the price-setting problem of producer i is

$$\max_{P_i} E_i\left[(P_i - \lambda_i)\, D_i\left(\left\{\frac{P_i}{P}\right\}_{i=0}^{1}, \{\gamma_i\}_{i=0}^{1}, Y\right)\right]. \tag{58}$$

The optimal price of intermediate good i is in this case

$$P_i = \frac{1}{z_i}\left(\frac{R}{\alpha}\right)^{\alpha}\left(\frac{W}{1-\alpha}\right)^{1-\alpha}\frac{E_i\Theta_i}{E_i\Theta_i - 1}, \tag{59}$$

where

$$\Theta_i = -\frac{\dfrac{\partial D_i\left(\left\{\frac{P_i}{P}\right\}_{i=0}^{1}, \{\gamma_i\}_{i=0}^{1}, Y\right)}{\partial \frac{P_i}{P}}\frac{P_i}{P}}{D_i\left(\left\{\frac{P_i}{P}\right\}_{i=0}^{1}, \{\gamma_i\}_{i=0}^{1}, Y\right)} \tag{60}$$

https://doi.org/10.1017/S1365100514000753 Published online by Cambridge University Press

is the elasticity of demand for intermediate good i. This elasticity is a function of all the relative prices $\{P_i/P\}_{i=0}^1$, all shocks $\{\gamma_i\}_{i=0}^1$, and the aggregate demand for goods Y, which are all unknown to producer i at the time it needs to set its price P_i. Hence, this framework illustrates that the earlier result, whereby producers' optimal price-setting (14) relies on just knowing their nominal costs and the exogenously given θ, is a special case that follows from the particular production function (1); it is not a general property.

Intermediate-good producer i could learn about its elasticity Θ_i by studying how the demand for its product changes as its relative price changes, computing

$$-\frac{\left(\frac{\Delta X_i}{\Delta \frac{P_i}{P}}\right) P_i/P}{X_i}, \tag{61}$$

where ΔX_i denotes the change in its demand, whereas $\Delta P_i/P$ is the change in its relative price. This would require knowing its relative price P_i/P, which requires knowing the aggregate price level P.[18] In addition, it would require that all other variables in the expression for Θ_i (60) remain constant over the period studied, as well as the demand function D_i itself, or that the producer control for the impact of such changes on the measured elasticity (61). In particular, this applies to aggregate output Y when the elasticity Θ_i varies over the business cycle, which it is likely to do [Domowitz et al. (1986a), Haskel et al. (1995), and Hall (2012) find that markups are procyclical in U.S. data, whereas Bils (1987), Galeotti and Schianterelli (1998), Rotemberg and Woodford (1999), and Gali et al. (2007) find that they are countercyclical]. Hence, even if price-setters simply mark up their nominal production costs, which we assume they can observe perfectly, their pricing decisions depend on aggregate variables they cannot immediately observe, such as the aggregate price and output levels, and therefore have to be based on their expectations of these. When such expectations are heterogeneous, as surveys show they are, relative prices are distorted, with some producers applying markups that are too high and others applying markups that are too low, affecting output through productivity.

By reducing the effectiveness with which labor and capital can be used to produce final goods, heterogeneous expectations have a negative impact on consumers' lifetime utility (20), just as a standard negative productivity shock would. As a result, there is scope for improving welfare by reducing such dispersion. Because heterogeneous expectations about aggregates can be one of its sources, there would be gains to homogenizing these. In particular, the aggregate price level lends itself to this, because according to the money-market clearing condition (36), it can be controlled through the money supply. The more credible a target the monetary authority can provide for the aggregate price level, or the rate at which it changes, the rate of inflation, the less dispersion is generated by the uncertainty inherent in its contemporary value, and the more efficient the production of final goods.[19] Although the dispersion is independent of the degree to which such a target is met, as long as it synchronizes expectations, the credibility of such a regime going forward, and therefore its ability to reduce future dispersion, require matching the target closely.

9. CONCLUSIONS

In a general equilibrium model with perfectly flexible prices, we show how heterogeneous expectations can reduce output by distorting relative prices and reducing productivity. Moreover, we find that shocks to the dispersion in expectations can be indistinguishable from standard productivity shocks and can be a significant source of business-cycle fluctuations.

NOTES

1. Sticky-price and sticky-information models introduce menu costs and informational delays, but otherwise assume that price-setters simply apply an exogenously given markup. In contrast to our model, distortions to relative prices arise because of the staggered adjustment of prices caused by the costs of updating information and changing prices. See Rotemberg (1982), Calvo (1983), and Mankiw and Reis (2002) for examples.
2. Expectations generating business cycles is not new; see, for example, Benhabib and Farmer (1994), Pintus (2011), Lasselle et al. (2005), Beaudry and Portier (2007), Eusepi and Preston (2008), and Nourry and Venditti (2012).
3. Mankiw et al. (2003) document the disparities that exist in inflation forecasts.
4. Expectations would become more homogeneous as price-setters obtained longer histories of their own observations, or if information were shared across firms.
5. The γ_i shocks could also be interpreted as taste shocks, but then the composition of the final good would change over time, making it difficult to compare final goods produced at different times.
6. Including land or any other inelastically supplied input facilitates obtaining an explicit solution for aggregate output. The reason is that the production side only pins down the optimal factor mix, so fixing the level of one of the inputs pins down all other levels. Land plays no other role in the model, and we let its importance in production (ν) converge to zero later.
7. We use the convention that it is capital K_{t-1} that is available to produce in period t, so that intermediate-good producer i is renting $k_{i,t-1}$ units of capital in period t.
8. According to the demand function (37), forecast errors $E_i\theta - \theta$ could be correlated with autocorrelated γ_i-shocks.
9. Modeling dispersion in terms of the markups directly, without any skewness, makes dispersion have a positive impact on the rental rate.
10. The income share of labor is usually taken to be between 0.6 and 0.7, with $\frac{2}{3}$ being the most popular value. We have chosen a value at the lower end because it allows the elasticity θ to be smaller without α going negative. With a labor share of 0.7, having a strictly positive value of α requires $\theta > 3.33$; with a share of 2/3 it requires $\theta > 3$. The value of α affects the steady-state effects of dispersion $\bar{r}$, but has no impact on the immediate effects $\bar{r}_t$ and $\bar{q}_t$.
11. We use 25 million producers in our simulations, but the results do not depend on this number as long as we have enough firms, or replications, for the simulated samples to be representative for the complete distribution.
12. When $\sigma_\gamma \leq 0.55$, this approximation is at most 1.31% off for $\theta \geq 7.5$ and less than 3.2% off for $\theta \geq 2.51$. It is more accurate the lower the value of σ_γ and the higher the value of θ. The approximation is imperfect because r_t drops faster than q_t when σ_γ increases, as is reflected in Figures 1 and 2.

13. Comparing Figures 1 and 2 shows that dispersion actually has a slightly larger immediate impact on output than on the rental rates, which makes it differ from the impact of a traditional productivity shock, Z_t. The approximation (51) between q_t and r_t eliminates this difference, which is quite small.

14. The values of b and μ have no effect on the impact dispersion r_t has on output Y_t, just as they have no effect on the impact the productivity shock Z_t has on Y_t. For $\sigma_\gamma \leq 0.55$, these effects are also independent of the calibrated value of σ_γ, which affects the steady-state value of r_t.

15. This number is consistent with what is commonly found for traditional productivity shocks in models with perfect competition; see for example Cooley (1995) and Walsh (2003). The impact is largest immediately, and gradually fades away.

16. With $\rho_Z = 0.95$, $\beta = 0.995$, $\delta = 0.019$, $\phi = 0.2$, $G/Y = 0.15$, $N = 0.2$, and a labor share of $\frac{2}{3}$, a 1% deviation in r_t from steady state makes output deviate 1.92–1.97% from steady state for $\theta \geq 4$. With $\rho_Z = 0.95$, $\beta = 0.96$, $\delta = 0.04$, $\phi = 0.9$, $G/Y = 0.4$, $N = 0.4$, and a labor share of 0.6, we get an output response of 1.00–1.06% for $\theta \geq 4$.

17. Allowing for variable capital utilization would raise the impact that a dispersion shock r_t has on output. However, making capital utilization endogenous would not change our results much for lower values of θ, because capital then carries a smaller weight in production. Allowing for variable effort in labor would have a greater impact, a situation that can be approximated by raising the intertemporal substitution, that is, lowering ϕ. With $\phi = 0.2$, the numbers for the impact of dispersion shocks presented previously increase only by 7–8% for $\theta \geq 3$ when $\rho_r = 0.95$. When $\rho_r = 0.5$ the numbers increase by 11–18% for $\theta \geq 3$, with the larger increases occurring for lower values of θ. When $\theta = 2.51$, changing ϕ has practically no impact.

18. Even if the aggregate price level P were constant, producers would need to know its level to know how much their relative price was changing. The assumption that intermediate-good producers do not know the aggregate price level, whereas final-good producers do, is a simplification. In reality, with heterogeneous households, each would know how the cost of its particular consumption bundle evolves, but not how that of households with dissimilar tastes behaves. A household's demand for a particular intermediate good then depends on its price relative to that of the household's preferred bundle, which will vary across households. Hence, intermediate good producers need to know the costs of all households' bundles, not just the aggregate price level. The same applies to other aggregate variables.

19. See Bernanke and Mishkin (1997), McCallum (1996), Svensson (1997 and 1999), and Bernanke and Woodford (2005) for discussions of inflation targeting. With lump-sum taxation, the welfare-maximizing inflation target in the model is the Friedman (1969) rule, that is, making the rate of inflation equal the negative of the real rate of return on capital.

REFERENCES

Beaudry, P. and R. Portier (2007) When can changes in expectations cause business cycle fluctuations in neo-classical settings? *Journal of Economic Theory* 135, 458–477.

Benhabib, J. and R. Farmer (1994) Indeterminacy and increasing returns. *Journal of Economic Theory* 63, 19–41.

Bernanke, B.S. and F.S. Mishkin (1997) Inflation targeting: A new framework for monetary policy? *Journal of Economic Perspectives* 11, 97–116.

Bernanke, B.S. and M. Woodford (2005) *The Inflation-Targeting Debate.* Chicago: University of Chicago Press.

Bils, M. (1987) The cyclical behavior of marginal cost and price. *American Economic Review* 77, 838–855.

Blanchard, O.J. and N. Kiyotaki (1987) Monopolistic competition and the effects of aggregate demand. *American Economic Review* 77, 647–666.

Bresnahan, T.F. and P.C. Reiss (1991) Entry and competition in concentrated markets. *Journal of Political Economy* 99, 977–1009.

Caballero, R.J., E.M.R.A. Engel, and J. Haltiwanger (1997) Aggregate employment dynamics: Building from microeconomic evidence. *American Economic Review* 87, 115–137.

Calvo, G.A. (1983) Staggered prices in a utility maximizing framework. *Journal of Monetary Economics* 12, 383–398.

Campbell, J. and H. Hopenhayn (2005) Market size matters. *Journal of Industry Studies* 53, 1–25.

Comin, D. and S. Mulani (2006) Diverging trends in aggregate and firm volatility. *Review of Economics and Statistics* 88, 374–383.

Cooley, T.F. (ed.) (1995) *Frontiers of Business of Business Cycle Research.* Princeton, NJ: Princeton University Press.

Cooley, T.F. and G.D. Hansen (1995) Money and the business cycle. In T.F. Cooley (ed.), *Frontiers of Business of Business Cycle Research*. Princeton, NJ: Princeton University Press.

Davis, S.J., J. Haltiwanger, and S. Schuh (1996) Small business and job creation: Dissecting the myth and reassessing the facts. *Small Business Economics* 8, 297–315.

Dixit, A.K. and J.E. Stiglitz (1977) Monopolistic competition and optimum product diversity. *American Economic Review* 67, 297–308.

Domowitz, I., R.G. Hubbard, and B.C. Petersen (1986a) Business cycles and the relationship between concentration and price–cost margins. *Rand Journal of Economics* 17, 1–17.

Domowitz, I., R.G. Hubbard, and B.C. Petersen (1986b) The intertemporal stability of the concentration–margins relationship. *Journal of Industrial Economics* 35, 13–34.

Eisfeldt, A.L. and A.A. Rampini (2006) Capital reallocation and liquidity. *Journal of Monetary Economics* 53, 369–399.

Eusepi, S. and B. Preston (2011) Expectations, learning and business cycle fluctuations. *American Economic Review* 101, 2844–2872.

Friedman, M. (1969) The optimum quantity of money. In M. Friedman (ed.), *The Optimum Quantity of Money and Other Essays*, pp. 1–50. Chicago: Aldine Publishing.

Gabaix, X. (2011) The granular origins of aggregate fluctuations. *Econometrica* 79, 733–772.

Galeotti, M. and F. Schianterelli (1998) The cyclicality of markups in a model with adjustment costs: Econometric evidence for US industry. *Oxford Bulletin of Economics and Statistics* 60, 121–142.

Gali, J., M. Gertler, and J.D. Lopez-Salido (2007) Markups, gaps, and the welfare costs of business fluctuations. *Review of Economics and Statistics* 89, 44–59.

Hall, R.E. (2012) The Cyclical Response of Advertising Refutes Counter-cyclical Profit Margins in Favor of Product Market Frictions. NBER working paper 18370.

Hansen, G.D. (1997) Technical progress and aggregate fluctuations. *Journal of Economic Dynamics and Control* 21, 1005–1023.

Haskel, J., C. Martin, and I. Small (1995) Price, marginal cost and the business cycle. *Oxford Bulletin of Economics and Statistics* 57, 25–41.

Hodrick, R.J. and E.C. Prescott (1997) Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking* 29, 1–16.

Kehrig, M. (2011) The Cyclicality of Productivity Dispersion. Working paper 11-15, Center for Economic Studies, U.S. Census Bureau.

Kydland, F.E. and E.C. Prescott (1982) Time to build and aggregate fluctuations. *Econometrica* 50, 1345–1370.

Lasselle, L., S. Svizzero, and C. Tisdell (2005) Stability and cycles in a cobweb model with heterogeneous expectations. *Macroeconomic Dynamics* 9, 630–650.

Lerner, A.P. (1934) The concept of monopoly and the measurement of monopoly power. *Review of Economic Studies* 1, 157–175.

Lucas, R.E. (1972) Expectations and the neutrality of money. *Journal of Economic Theory* 4, 103-124.

Mankiw, N.G. and R. Reis (2002) Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve. *Quarterly Journal of Economics* 117, 1295–1328.

Mankiw, N.G., R. Reis, and J. Wolfers (2003) Disagreement about Inflation Expectations. NBER working paper 9796.

Martins, J., S. Scapetta, and D. Pilat (1996) Markup pricing, market structure and the business cycle. *OECD Economic Studies* 27, 71–105.
McCallum, B.T. (1996) Inflation Targeting in Canada, New Zealand, Sweden, the United Kingdom, and in General. NBER working paper 5579.
Nourry, C. and A. Venditti (2012) Endogenous business cycles in overlapping-generations economies with multiple consumption goods. *Macroeconomic Dynamics* 16, 86–102.
Pintus, P.A. (2004) Expectations-driven fluctuations when factor utilization is variable. *Macroeconomic Dynamics* 8, 3–26.
Roberts, M.J. and D. Supina (2000) Output price and markup dispersion in producer micro data: The roles of producer heterogeneity and noise. *Industrial Organization* 9, 1–36.
Rotemberg, J.J. (1982) Monopolistic price adjustment and aggregate output. *Review of Economic Studies* 44, 254–281.
Rotemberg, J.J. and M. Woodford (1999) The cyclical behavior of prices and costs. In J.B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, pp. 1051–1135. Amsterdam: Elsevier.
Svensson, L.E.O. (1997) Optimal inflation targets, “conservative” central banks, and linear inflation contracts. *Journal of Monetary Economics* 50, 691–720.
Svensson, L.E.O. (1999) Inflation targeting as a monetary policy rule. *Journal of Monetary Economics* 43, 607–654.
Walsh, C.E. (2003) *Monetary Theory and Policy*. Cambridge, MA: MIT Press.