Logit price dynamics

James Costain* and Anton Nakov†
*Banco de España
†Federal Reserve Board

May 2012
PRELIMINARY AND INCOMPLETE

Abstract

We propose a tractable near-rational model of retail price adjustment consistent with microeconomic and macroeconomic evidence on price dynamics. Our framework assumes firms may make errors either in the *size* or the *timing* of their price adjustments. Both types of pricing mistakes are governed by a logit distribution. We show how this behavior can be built into a DSGE model and simulated using a straightforward algorithm for heterogeneous-agent computations.

The special case of our model with errors in the size of the adjustments but not in their timing is identical to the "precautionary price stickiness" setup of Costain and Nakov (2011C). This version of the model is consistent with several "puzzles" in microeconomic data on price adjustment: (1) large and small price adjustments coexist; (2) the price adjustment hazard declines with the time since last adjustment; (3) the size of the adjustment is largely independent of the time since last adjustment; and (4) extreme prices are more likely to be new than prices near the center of the distribution. The special case of our model with errors in the timing of adjustment but not in their size implies the same binary logit derived by Woodford (2009). This case implies a high degree of monetary nonneutrality, as macroeconomic studies of monetary policy shocks find. The general model with both types of errors, when parameterized to reproduce the histogram of price adjustments and the mean adjustment frequency, matches both the microeconomic and macroeconomic facts well.

Our logit assumption can be justified by appealing to a cost function for error avoidance if, as in Stahl (1990) or Mattsson and Weibull (2002), control costs are proportional to entropy reduction. Thus, this interpretation of our model implies price adjustment is costly, but the costs are not "menu costs". Instead, the costs are related to managerial decision-making, consistent with the evidence of Zbaracki *et al.* (2004). Costly choice of the *timing* of price adjustment implies behavior like a stochastic menu cost model (and offers an alternative justification for the adjustment hazard derived by Woodford, 2009) but, unlike those models, does not require the implausible assumption of *i.i.d.* adjustment costs.

A major technical advantage of our framework is that the firm's idiosyncratic state variable is simply its price level and productivity, whereas in Sims (2003) and the "rational inattention" literature the firm's idiosyncratic state is its *prior* (which is generally an infinite-dimensional object).

Keywords: Nominal rigidity, logit equilibrium, entropic equilibrium, state-dependent pricing, near rationality, information-constrained pricing

JEL Codes: E31, D81, C72

1 Introduction[1]

Economists seeking to explain price stickiness have often appealed to small fixed costs of nominal price changes, commonly called "menu costs" (Barro 1972). If shocks to fundamentals accumulate relatively slowly, then even small menu costs might suffice to make price adjustments infrequent and to make aggregate dynamics deviate in a nontrivial way from the flexible-price optimum (Mankiw 1985). However, Golosov and Lucas (2007) showed quantitatively that in a standard DSGE framework, fixed menu costs do little to generate aggregate price stickiness. The dynamics of their model are quite close to monetary neutrality, so fixed menu costs do not seem promising to explain the substantial real effects of monetary shocks observed in macroeconomic data (*e.g.* Christiano, Eichenbaum, and Evans, 2008). Moreover, detailed microeconomic evidence suggests that fixed costs, as usually interpreted, are only a small fraction of the overall costs of price setting (Zbaracki *et al.* 2004). A much larger part of the costs of price adjustment consists of managerial costs associated with information collection and decision making. This raises the question: does costly decision making explain microeconomic and macroeconomic evidence of price stickiness better than fixed menu costs do? And, at a more practical level, how exactly should costly price setting decisions be modeled?

This paper proposes a tractable way of modeling price stickiness, based on costly decision-making, and shows that it is consistent with a variety of microeconomic facts about price adjustments, as well as macroeconomic evidence of monetary nonneutrality. More precisely, we consider *two* game-theoretic foundations for price stickiness. On one hand, if decision-making is costly, then choices will typically be imperfect. That is, choices will be subject to errors, making it natural to think of the outcomes of decisions as random variables, rather than treating the choice made as a deterministic function of fundamentals. As is common in experimental economics and econometrics, we allow for errors by assuming that the distribution of choices is described by a logit. On the other hand, if choices are costly, it is natural to assume that more precise decisions are more costly than imprecise ones. Therefore we also consider a setup in which firms face a cost function that increases with the precision of their decisions.

As it happens, these two ways of modeling price stickiness are closely related. If the cost of decision-making increases linearly with precision, and precision is measured by relative entropy, then the optimal distribution for the decision variable is a logit. This is a well-known mathematical fact that was first discovered in the context of thermodynamics, where a formally identical problem gives rise to the Boltzmann distribution of particles in a gas. The fact that a logit distribution can be derived from a cost function based on entropy has been pointed out repeatedly in the game theory and economics literature (Stahl 1990; Marsili 1999; Mattson and Weibull 2002; Wolpert???, Matejka and McKay 2011). However, economics applications have typically focused on a single decision taken at a given point in time. Applying this insight to price adjustment requires us to consider errors on two recurrent margins of decision: *which* price the firm sets, and *when* the price is adjusted. We show how to extend the deriviation of a logit distribution of decisions from an entropy-related cost function to a fully dynamic context. In order to ensure that the empirical implications of the model are invariant to the assumed time period, the decision of whether or not to adjust at a given point must be treated as a *weighted*

[1]We thank Bartosz Mackowiak, Filip Matejka, Galo Nuño, Antonella Tutino, Mirko Wiederholt, Carl Walsh, Hervé Le Bihan, and seminar participants at UC Santa Cruz and T2M (Nantes) for helpful comments. Views expressed here are those of the authors and do not necessarily coincide with those of the Bank of Spain, the Eurosystem, or the Federal Reserve.

binary logit. In other words, while a standard static logit model has a single free parameter representing the accuracy of decisions, the weighted logit in our dynamic setup has two free parameters, representing the *speed* and the *accuracy* of decision making.

Summarizing our main findings, logit equilibrium fits micro price adjustment data well, in spite of its small number of free parameters. Our model performs well in terms of several of the stylized facts we just mentioned. It implies that many large and small price changes coexist, in contrast to the implications of a fixed menu cost model (Klenow and Kryvstov 2008; Midrigan 2011; Klenow and Malin 2010, "Fact 7"). It implies that the probability of price adjustment is nearly flat, but slightly decreasing in the first few months. This is the pattern found in empirical studies after allowing for heterogeneity in adjustment frequency (Nakamura and Steinsson 2008, "Fact 5"; Klenow and Malin 2010, "Fact 10"). Furthermore, we find that the standard deviation of price adjustment is mostly constant, independent of the time since last adjustment (Klenow and Malin 2010, "Fact 10"). Most alternative models, including the Calvo model, instead imply that price adjustments are increasing. Finally, our model implies that extremely high or low prices are more likely to have been set recently than prices near the center of the distribution (Campbell and Eden 2010).

Finally, we calculate the effects of trend inflation and money supply shocks in our framework. Given the degree of rationality that best fits microdata, the effect of money shocks on consumption is large, though smaller than the degree of nonneutrality found in the Calvo model. Nonneutrality results mainly from mistakes in the timing of price adjustments. While the version of the model with mistakes in the *size* of adjustment only helps qualitatively to match some microeconomic facts, quantitatively matching the both the size and frequency of price adjustments requires the version of the model with mistakes both in size and timing. The model with both types of mistakes has only two free parameters, but is nonetheless quite successful in reproducing steady-state price adjustment behavior, and in modeling how price adjustments vary with large changes in underlying trend inflation, and in obtaining substantial monetary nonneutrality in response to money growth shocks.

1.1 Related literature

This paper has links to several diverse areas of economic literature. It is most directly related to recent papers on state dependent pricing driven by aggregate and firm-specific shocks, including Golosov and Lucas (2007), Midrigan (2011), Dotsey, King, and Wolman (2008), Álvarez, Lippi, and Paciello (2011), Álvarez, Beraja, Gonzalez, and Neumeyer (2011), Kehoe and Midrigan (2010), and Matejka (2011). It builds specifically on two previous papers of our own: in Costain and Nakov (2011C) we study the microeconomic and macroeconomic implications of logit errors in price decisions, while one specification considered in Costain and Nakov (2011B) imposes logit errors on the timing of price adjustment. All this recent work on state-dependent pricing has been motivated by empirical studies of new microeconomic data on price adjustment derived from scanners used in retail scales, including Klenow and Kryvtsov (2008), Nakamura and Steinsson (2008), Klenow and Malin (2010), and Eichenbaum, Jaimovich, and Rebelo (2011).

While recent papers on state-dependent pricing mostly assume prices are set optimally subject to menu costs, this paper instead assumes price setting is subject to errors, and it does not assume any menu costs, at least not as they are usually interpreted. Our framework for modeling error-prone behavior, logit equilibrium, has been widely applied in experimental game theory, where it has helped explain play in a number of games where Nash equilibrium performs poorly, such as the centipede game and Bertrand competition games (McKelvey and Palfrey 1998; Anderson, Goeree, and Holt 2002). It has been much less frequently applied in other areas of economics; we are unaware of any application of logit equilibrium inside a dynamic general equilibrium macroeconomic model, other than our own work.[2] The absence of logit modeling in macroeconomics may be due, in part, to discomfort with the many degrees of freedom opened up by moving away from the benchmark of full rationality. However, since logit equilibrium is just a one-parameter or two parameter generalization of fully rational choice, it actually imposes much of the discipline of rationality on the model.[3]

McKelvey and Palfrey initially defined logit equilibrium both for normal form (1995) and extensive form (1998) games. However, it is by no means obvious how to apply the fixed sequence of choice times assumed in an extensive-form game to micro pricing data where the key issue under study is *when* adjustment occurs. Applying logit equilibrium concepts to the timing of adjustment requires us to search for a specification with empirical implications that are invariant to the choice of the frequency at which the model is studied. The correct way to do this becomes especially clear when we derive logit equilibrium from entropy control costs. In a static model, logit choice is derived by penalizing the relative entropy of the choice process, relative to a uniform distribution. Likewise, we derive a weighted binary logit that determines when adjustment occurs by penalizing the relative entropy of the random time of adjustment, relative to a constant probability of adjustment. In other words, precision in the *size* of the adjustment is measured by comparing the price distribution to a uniform distribution; likewise, precision in the *timing* of adjustment is measured by comparing the state-dependent hazard rate *to a Calvo model*. The hazard rate we derive from this specification has exactly the same functional form derived by Woodford (2009), though his microfoundations differ (he assumes a fixed cost of purchasing information and a fixed cost to adjust the price).

Woodford (2009) is one of several recent papers on "rational inattention" that have followed Sims (2003) by assuming economic agents face costs or constraints associated with information flow in the sense of Shannon (1948). Since we use changes in entropy to measure the precision of choices, and Shannon (1948) uses changes in entropy to measure information flow, our paper shares some of the mathematics in Sims (2003). The only difference is that in our setup, the probability distribution over a firm's decisions is conditioned on the firm's true state, whereas in the rational inattention literature the distribution of decisions is conditional on the firm's prior about its true state. In other words, Sims assumes the true state of the world is never known with certainty, so that receiving information forms part of the costs of decision-making. We instead simplify by assuming that the true state of the world is known, but that nonetheless making optimal choices is costly; that is, we focus on costs of information processing rather than costs of information reception. On the other hand, our model may be a more literally realistic description of price adjustment in a retail context. We consider a firm that intermittently sets a price, and model the costs of choosing when to adjust as well as the costs of choosing what price to set; this seems to be a fairly literal description of the variables actually controlled by many

[2]The logit choice function is probably the most standard econometric framework for discrete choice, and has been applied to a huge number of microeconometric contexts. But logit *equilibrium*, in which each player makes logit decisions, based on payoff values which depend on other players' logit decisions, has to the best of our knowledge rarely been applied outside of experimental game theory.

[3]Haile, Hortaçsu, and Kosanok (2008) have shown that quantal response equilibrium, which has an infinite number of free parameters, is impossible to reject empirically. However, this criticism does not apply to logit equilibrium (the special case of quantal response equilibrium which has been most widely applied in practice) since it is very tightly parameterized.

retail firms. Sims instead models prices as a process that may vary continuously in continuous time, which is a more metaphorical description of price setting.

But ultimately, our reason for proceding as we do is a practical one: our model is "infinitely" easier to solve than that of Sims (2003) and Matejka (2011). Our main reason for ignoring the first stage of the information flow is that by doing so we dramatically reduce the dimension of the calculations required to solve our model. Since the rational inattention approach assumes the firm acts under uncertainty, it implies the firm conditions on a *prior* over its possible productivity levels (which is a high-dimensional object that complicates solution of the model). In our setup, the firm just conditions on its true productivity level. Moreover, once one knows that entropy reduction costs imply logit, one can simply impose a logit function directly (and then subtract off the implied costs) rather than explicitly solving for the form of the error distribution. These facts make our approach entirely tractable in a DSGE context, as this paper will show.

2 Model

This discrete-time model embeds near-rational price adjustment by firms in an otherwise standard New Keynesian general equilibrium framework based on GL07. Besides the firms, there is a representative household and a monetary authority that either implements a Taylor rule or follows an exogenous growth process for nominal money balances.

The aggregate state of the economy at time t, which will be identified in Section 2.3, is called Ω_t. Whenever aggregate variables are subscripted by t, this is an abbreviation indicating dependence, in equilibrium, on aggregate conditions Ω_t. For example, consumption is denoted by $C_t \equiv C(\Omega_t)$.

2.1 Household

The household's period utility function is $\frac{1}{1-\gamma}C_t^{1-\gamma} - \chi N_t + \nu \log(M_t/P_t)$, where C_t is consumption, N_t is labor supply, and M_t/P_t is real money balances. Utility is discounted by factor β per period. Consumption is a CES aggregate of differentiated products C_{it}, with elasticity of substitution ϵ:

$$C_t = \left\{ \int_0^1 C_{it}^{\frac{\epsilon-1}{\epsilon}} di \right\}^{\frac{\epsilon}{\epsilon-1}}. \tag{1}$$

The household's nominal period budget constraint is

$$\int_0^1 P_{it}C_{it}di + M_t + R_t^{-1}B_t = W_tN_t + M_{t-1} + T_t + B_{t-1} + Z_t \tag{2}$$

where $\int_0^1 P_{it}C_{it}di$ is total nominal consumption. B_t represents nominal bond holdings, with interest rate $R_t - 1$; T_t is a lump sum transfer from the central bank, and Z_t is a dividend payment from the firms.

Households choose $\{C_{it}, N_t, B_t, M_t\}_{t=0}^{\infty}$ to maximize expected discounted utility, subject to the budget constraint (2). Optimal consumption across the differentiated goods implies

$$C_{it} = (P_t/P_{it})^{\epsilon}C_t, \tag{3}$$

so nominal spending can be written as $P_t C_t = \int_0^1 P_{it} C_{it} di$ under the price index

$$P_t \equiv \left\{ \int_0^1 P_{it}{}^{1-\epsilon} di \right\}^{\frac{1}{1-\epsilon}}. \tag{4}$$

Defining inflation as $\Pi_{t+1} \equiv P_{t+1}/P_t$, the first-order conditions for labor supply, consumption, and money use can be written as:

$$\chi = C_t^{-\gamma} W_t / P_t, \tag{5}$$

$$R_t^{-1} = \beta E_t \left(\frac{C_{t+1}^{-\gamma}}{\Pi_{t+1} C_t^{-\gamma}} \right), \tag{6}$$

$$1 - \frac{v'(m_t)}{C_t^{-\gamma}} = \beta E_t \left(\frac{C_{t+1}^{-\gamma}}{\Pi_{t+1} C_t^{-\gamma}} \right). \tag{7}$$

2.2 Monopolistic firms

Each firm i produces output Y_{it} under a constant returns technology $Y_{it} = A_{it} N_{it}$, where A_{it} is an idiosyncratic productivity process, AR(1) in logs:

$$\log A_{it} = \rho \log A_{it-1} + \varepsilon_{it}^a, \tag{8}$$

and labor N_{it} is the only input. Firm i is a monopolistic competitor that sets a price P_{it}, facing the demand curve $Y_{it} = C_t P_t^{\epsilon} P_{it}^{-\epsilon}$, and must fulfill all demand at its chosen price. It hires in a competitive labor markets at wage rate W_t, generating profits

$$U_{it} = P_{it} Y_{it} - W_t N_{it} = \left(P_{it} - \frac{W_t}{A_{it}} \right) C_t P_t^{\epsilon} P_{it}^{-\epsilon} \equiv U(P_{it}, A_{it}, \Omega_t) \tag{9}$$

per period. Firms are owned by the household, so they discount nominal income between times t and $t+1$ at the rate $\beta \frac{P(\Omega_t) u'(C(\Omega_{t+1}))}{P(\Omega_{t+1}) u'(C(\Omega_t))}$, consistent with the household's marginal rate of substitution.

2.2.1 The size of price adjustments

Let $V(P_{it}, A_{it}, \Omega_t)$ denote the nominal value of a firm at time t that produces with productivity A_{it} and sells at nominal price P_{it}. Since we assume firms are not costlessly capable making of precisely optimal choices, the nominal price P_{it} at a given point in time will not necessarily be optimal. Indeed, we assume decisions are subject to errors, meaning that the firm's price-setting process will determine a *conditional distribution* $\pi(P|A_{it}, \Omega_t)$ across possible prices, rather than picking out a single optimal value. The key property we impose on the distribution π is that the probability of choosing any given price is a smoothly increasing function of the value of choosing that price.

As is common in microeconometrics and experimental game theory, we assume the distribution of errors is given by a multinomial logit. In order to treat the logit function as a primitive of the model, we define its argument in units of labor time. That is, since the costs of decision-making are presumably related to the labor effort (in particular, managerial labor) required to calculate and communicate the chosen price, we divide the values in the logit function by the wage rate, $W(\Omega_t)$, to convert them to time units. Hence, the probability $\pi(P^j|A_{it},\Omega_t)$ of choosing price $P^j \in \Gamma^P$ at time t, conditional on productivity A_{it}, is given by

$$\pi(P^j|A_{it},\Omega_t) \equiv \frac{\exp\left(\frac{V(P^j,A_{it},\Omega_t)}{\kappa W(\Omega_t)}\right)}{\sum_{k=1}^{\#P}\exp\left(\frac{V(P^k,A_{it},\Omega_t)}{\kappa W(\Omega_t)}\right)} \tag{10}$$

Note that for numerical purposes, we constrain the price choice to a finite discrete grid $\Gamma^P \equiv \left\{P^1, P^2, ...P^{\#P}\right\}$. The parameter κ in the logit function can be interpreted as the degree of noise in the decision process; in the limit as $\kappa \to 0$ it converges to the policy function under full rationality, so that the optimal price is chosen with probability one.[4]

We will use the notation E^π to indicate an expectation taken under the logit probability (10). The firm's expected value, conditional on adjusting to a new price $P' \in \Gamma^P$, is then

$$E^\pi V(P',A_{it},\Omega_t) \equiv \sum_{j=1}^{\#P}\pi(P^j|A_{it},\Omega_t)V(P^j,A_{it},\Omega_t) \tag{11}$$

$$= \sum_{j=1}^{\#P}\frac{\exp\left(\frac{V(P^j,A_{it},\Omega_t)}{\kappa W(\Omega_t)}\right)V(P^j,A_{it},\Omega_t)}{\sum_{k=1}^{\#P}\exp\left(\frac{V(P^k,A_{it},\Omega_t)}{\kappa W(\Omega_t)}\right)} \tag{12}$$

Given the potential for errors, it may or may not be profitable for the firm to adjust to a new price. For clarity, it helps to distinguish the firm's beginning-of-period price, $\widetilde{P}_{it} \equiv P_{i,t-1}$, from the end-of-period price P_{it} that its time t customers pay, which may or may not be the same. For a firm that begins period t with price $\widetilde{P}_{it}$, the gain from adjusting at the beginning of t is:

$$D(\widetilde{P}_{it},A_{it},\Omega_t) \equiv E^\pi V(P',A_{it},\Omega_t) - V(\widetilde{P}_{it},A_{it},\Omega_t). \tag{13}$$

Evidently, if $\widetilde{P}_{it}$ is already close to the optimal price $P^*(A_{it},\Omega_t) \equiv \text{argmax}_P V(P,A_{it},\Omega_t)$, then $D(\widetilde{P}_{it},A_{it},\Omega_t)$ may be negative, implying that it is better to avoid the risk of price-setting errors by maintaining the current price.

2.2.2 The timing of price adjustments

Thus, the firm faces a binary choice at each point in time: should it adjust its price? Here again, we assume decisions are error-prone, and we impose a regularity condition analogous to the one we imposed before: the probability of price adjustment is a smoothly increasing function λ of the gain from adjustment. In order to take λ as a primitive of the model, we scale by the wage so that the argument of the function represents units of labor time. Thus, the probability of adjustment will be defined as $\lambda\left(L\left(\widetilde{P}_{it},A_{it},\Omega_t\right)\right)$, where $L\left(\widetilde{P}_{it},A_{it},\Omega_t\right) = \frac{D(\widetilde{P}_{it},A_{it},\Omega_t)}{W(\Omega_t)}$ expresses the gains from adjusting in time units by dividing by the wage.

The next question is what functional form to impose on λ. It might seem natural to impose a simple binary logit that compares the values of adjusting and not adjusting:

$$\frac{\exp\left(\frac{E^\pi V(P',A_{it},\Omega_t)}{\kappa W(\Omega_t)}\right)}{\exp\left(\frac{E^\pi V(P',A_{it},\Omega_t)}{\kappa W(\Omega_t)}\right)+\exp\left(\frac{V(P^j,A_{it},\Omega_t)}{\kappa W(\Omega_t)}\right)} = \left(1+\exp\left(\frac{-D(P^j,A_{it},\Omega_t)}{\kappa W(\Omega_t)}\right)\right)^{-1} \tag{14}$$

[4] Alternatively, logit models are often written in terms of the inverse parameter $\xi \equiv \kappa^{-1}$, which can be interpreted as a measure of the degree of rationality.

This smooth function (i) approaches 1 in the limit as the adjustment gain D increases, (ii) approaches 0 as $D \to -\infty$, and (iii) implies that if the firm is indifferent between adjusting and not adjusting ($E_t^\pi V(P', A_{it}, \Omega_t) = V(\widetilde{P}_{it}, A_{it}, \Omega_t)$) then the probability of adjustment in period t is 0.5. But upon reflection, (iii) cannot be a desirable property, because the period length imposed on the model is arbitrary.[5] But under functional form (14), the probability of adjustment conditional on indifference is one-half *regardless of the value of* κ and regardless of period length. Thus, for example, solving the model with adjustment probability (14) at weekly frequency would imply continuous-time adjustment rates *roughly four times as high* as an otherwise identical solution at monthly frequency.

This problem is solved if we instead impose a weighted binary logit, as follows:

$$\lambda\left(\frac{D(P^j, A_{it}, \Omega_t)}{\kappa W(\Omega_t)}\right) = \frac{\bar{\lambda}\exp\left(\frac{E_t^\pi V(P', A_{it}, \Omega_t)}{\kappa W(\Omega_t)}\right)}{\bar{\lambda}\exp\left(\frac{E_t^\pi V(P', A_{it}, \Omega_t)}{\kappa W(\Omega_t)}\right) + (1-\bar{\lambda})\exp\left(\frac{V(P^j, A_{it}, \Omega_t)}{\kappa W(\Omega_t)}\right)} \tag{15}$$

$$= \left(1 + \rho \exp\left(\frac{-D(P^j, A_{it}, \Omega_t)}{\kappa W(\Omega_t)}\right)\right)^{-1}. \tag{16}$$

where $\rho = (1-\bar{\lambda})/\bar{\lambda}$. Like (14), this weighted logit goes smoothly from 0 to 1 as the adjustment gain D goes from $-\infty$ to ∞. But when the firm is indifferent between adjusting and not adjusting ($D = 0$), the probability of adjustment is $\bar{\lambda}$, which is a free parameter. Just as κ^{-1} is related to the *accuracy* of price adjustment, $\bar{\lambda}$ is related to the *speed* of price adjustment. This additional free parameter can be scaled up or down so that the model can be defined at any (sufficiently short) discrete time period.[6]

2.2.3 The firm's Bellman equation

We are now ready to write a Bellman equation for the monopolistic competitor. The value of selling at any given price equals current profits plus the expected value of future production, which may or may not occur at a new, adjusted price. Given the firm's idiosyncratic state variables (P, A) and the aggregate state Ω, and denoting next period's variables with primes, the Bellman equation is

$$V(P, A, \Omega) = \left(P - \frac{W(\Omega)}{A}\right) C(\Omega) P(\Omega)^\epsilon P^{-\epsilon} + \tag{17}$$

$$\beta E\left\{ \frac{P(\Omega)C(\Omega')^{-\gamma}}{P(\Omega')C(\Omega)^{-\gamma}} \left[\left(1 - \lambda\left(\frac{D(P,A',\Omega')}{W(\Omega')}\right)\right) V(P, A', \Omega') + \lambda\left(\frac{D(P,A',\Omega')}{W(\Omega')}\right) E^\pi V(P', A', \Omega') \right] \Big| A, \Omega\right\}.$$

Here the expectation E refers to the distribution of A' and Ω' conditional on A and Ω, and E^π represents an expectation over P' conditional on (A', Ω'), as defined in (12). Note that on the left-hand side of the Bellman equation, and in the term that represents current profits, P refers to a given firm i's price P_{it} at the end of t, when transactions occur. In the expectation on the

[5] Model properties should be approximately invariant to period length as long as we choose a period sufficiently short so that the probability of adjusting *more than once* per period is relatively small over all states (P, A, Ω) that occur with nonnegligible probability in equilibrium.

[6] In principle, we could allow for yet another free parameter by allowing the noise parameter κ to differ the two logits π and λ. But when we derive the logits from control costs in Section 3 we will see that the two noise parameters should be the same, so we impose this restriction here too.

right, P represents the price $\widetilde{P}_{i,t+1}$ at the beginning of $t+1$, which may (probability λ) or may not $(1-\lambda)$ be adjusted prior to time $t+1$ transactions to a new value P'.

It may sound strange to hear (17) called a "Bellman equation" when it contains no "max" or "min" operator. But a certain degree of optimization is implicit in the probabilities π and λ: as $\kappa \to 0$, (17) places probability one on the optimal choice at each decision step, so it becomes a Bellman equation in the usual sense. More generally, (17) allows for errors, but it always places higher probability on better choices, except in the limit $\kappa \to \infty$, in which decisions are perfectly random.

The right-hand side of the Bellman equation can be simplified by using the notation from (9), and the rearrangement $(1-\lambda)V + \lambda E^{\pi}V = V + \lambda(E^{\pi}V - V)$:

$$V(P,A,\Omega) \;=\; U(P,A,\Omega) + \beta E\left\{ \tfrac{P(\Omega)C(\Omega')^{-\gamma}}{P(\Omega')C(\Omega)^{-\gamma}} \left[V(P,A',\Omega') + G(P,A',\Omega')\right] \Big|\, A,\Omega \right\}, \qquad (18)$$

where

$$G(P,A',\Omega') \equiv \lambda\left(\frac{D(P,A',\Omega')}{W(\Omega')}\right) D(P,A',\Omega'). \qquad (19)$$

The terms inside the expectation in the Bellman equation represent the value V of continuing without adjustment, plus the flow of expected gains G due to adjustment. Since the firm plays the logit (10) whenever it adjusts, the price process associated with (18) is

$$P_{it} = \begin{cases} P^j \in \Gamma^P & \text{with probability} \quad \lambda\left(\frac{D(\widetilde{P}_{it},A_{it},\Omega_t)}{W(\Omega_t)}\right)\pi(P^j|A_{it},\Omega_t) \\ \widetilde{P}_{it} \equiv P_{i,t-1} & \text{with probability} \quad 1-\lambda\left(\frac{D(\widetilde{P}_{it},A_{it},\Omega_t)}{W(\Omega_t)}\right) \end{cases}. \qquad (20)$$

Equation (20) is written with time subscripts for additional clarity.

2.2.4 Extreme special cases

This setup nests two special cases which we will compare with the general case in the simulations that follow. On one hand, we could allow for mistakes in the size of price adjustments, but assume that the timing of price adjustment is perfectly optimal. That is, we could assume that price resetting behavior is governed by the distribution (10), while the timing of resets is given by

$$\lambda(L) \;=\; \mathbf{1}(L \geq 0), \qquad (21)$$

so that adjustment occurs if and only if it increases value. Since the potential for errors in (10) makes price adjustment risky, it means firms will avoid adjusting whenever they are sufficiently close to the optimum, which is why we have called this specification "precautionary price stickiness" in an earlier paper.

At the opposite extreme, we could assume that any adjusting firm always sets the optimal price ($\pi_t(P^*,A) = 1$ if $P^* = \text{argmax}_P V(P,A)$, with probability zero for all other prices), while allowing for "mistakes" in the timing of price adjustment by imposing the weighted logit (16). Such a framework exhibits near-rational price stickiness, in the sense of Akerlof and Yellen (1985), since the probability of price adjustment increases smoothly with the value of adjustment, so firms frequently leave the price unchanged when the value of adjustment is small. We will call the functional form (16) for the adjustment probability "Woodford's logit", because Woodford (2009) derived it as a consequence of a Shannon constraint on information flow together with a fixed cost of purchasing information plus a fixed cost of price adjustment. In the next section, we will show that it can also be derived from costly error avoidance in the absence of any menu cost or other physical fixed costs.[7]

2.3 Monetary policy and aggregate consistency

Two specifications for monetary policy are compared: a money growth rule and a Taylor rule. In both cases the systematic component of monetary policy is perturbed by an AR(1) process z,

$$z_t = \phi_z z_{t-1} + \epsilon_t^z, \tag{22}$$

where $0 \leq \phi_z < 1$ and $\epsilon_t^z \sim i.i.d.N(0, \sigma_z^2)$. Under the money growth rule, which is analyzed first to build intuition and for comparison with previous studies, z affects money supply growth:

$$M_t/M_{t-1} \equiv \mu_t = \mu^* \exp(z_t). \tag{23}$$

Alternatively, under a Taylor interest rate rule, which is a better approximation to actual monetary policy, the nominal interest rate follows

$$\frac{R_t}{R^*} = \exp(-z_t) \left(\left(\frac{P_t/P_{t-1}}{\Pi^*} \right)^{\phi_\pi} \left(\frac{C_t}{C^*} \right)^{\phi_c} \right)^{1-\phi_R} \left(\frac{R_{t-1}}{R^*} \right)^{\phi_R}, \tag{24}$$

where $\phi_c \geq 0$, $\phi_\pi > 1$, and $0 < \phi_R < 1$, so that when inflation Π_t exceeds its target Π^* or consumption C_t exceeds its target C^*, R_t tends to rise above its target $R^* \equiv \Pi^*/\beta$. For comparability between the two monetary regimes, the inflation target is set to $\Pi^* \equiv \mu^*$, and the rules are specified so that in both cases, a positive z represents an expansive shock.

Seigniorage revenues are paid to the household as a lump sum transfer T_t, and the government budget is balanced each period, so that $M_t = M_{t-1} + T_t$. Bond market clearing is simply $B_t = 0$. When supply equals demand for each good i, total labor supply and demand satisfy

$$N_t = \int_0^1 \frac{C_{it}}{A_{it}} di = P_t^\epsilon C_t \int_0^1 P_{it}^{-\epsilon} A_{it}^{-1} di \equiv \Delta_t C_t. \tag{25}$$

Equation (25) also defines a measure of price dispersion, $\Delta_t \equiv P_t^\epsilon \int_0^1 P_{it}^{-\epsilon} A_{it}^{-1} di$, weighted to allow for heterogeneous productivity. As in Yun (2005), an increase in Δ_t decreases the goods produced per unit of labor, effectively acting like a negative aggregate shock.

At this point, all equilibrium conditions have been spelled out, so an appropriate aggregate state variable Ω_t can be identified. At time t, the lagged distribution of transaction prices $\Phi_{t-1}(P, A)$ is predetermined. Knowing Φ_{t-1}, the lagged price level can be substituted out of the Taylor rule, using $P_{t-1} = \left[\iint P^{1-\epsilon} \Phi_{t-1}(dP, dA) \right]^{1/(1-\epsilon)}$. It can thus be seen that $\Omega \equiv (z_t, R_{t-1}, \Phi_{t-1})$ suffices to define the aggregate state. Given this Ω_t, equations (4), (5), (6), (8), (9), (**??**), (18), (19), (20), (22), (24), and (25) together give enough conditions to determine the distributions $\widetilde{\Phi}_t$ and Φ_t, the price level P_t, the functions $V_t \equiv V(P, A, \Omega_t)$, U_t,

[7] Although we assume Woodford's functional form for the adjustment probability, this special case of our model is not exactly the same as Woodford (2009). Since he considered a rational inattention framework, the gains from adjustment in his model are evaluated in terms of a prior over possible values of the current state, whereas in our model the gains from adjustment are evaluatd in terms of the firm's true state.

D_t, and G_t, and the variables R_t, C_t, N_t, W_t, and z_{t+1}. Thus they determine the next state, $\Omega_{t+1} \equiv (z_{t+1}, R_t, \Phi_t)$.

Under a money growth rule, the time t state can instead be defined as $\Omega_t \equiv (z_t, M_{t-1}, \Phi_{t-1})$. Substituting (7) for (6) and (23) for (24), knowing $\Omega_t \equiv (z_t, M_{t-1}, \Phi_{t-1})$ suffices to determine $\widetilde{\Phi}_t$, Φ_t, P_t, V_t, U_t, D_t, G_t, C_t, N_t, W_t, z_{t+1}, and M_t. Thus the next state, $\Omega_{t+1} \equiv (z_{t+1}, M_t, \Phi_t(P, A))$, can be calculated.

3 Model: control costs

3.1 Discussion

3.2 Choosing a new price

Our logit assumption (10) has the desirable property that the probability of choosing any given price is a smoothly increasing function of the value of that price. We now show that the logit functional form can be derived from an assumption that precise managerial decisions are costly. That is, suppose that firms must pay "control costs",[8] defined in units of time, to make a more precise choice (equivalently, to decrease the error in their choice). We will follow Stahl (1990) and Mattsson and Weibull (2002) by assuming that the cost of increased precision is proportional to the reduction in the entropy of the choice variable, normalizing the cost of a perfectly random decision (a uniform distribution) to zero.[9]

This definition of the cost function can also be stated in terms of the statistical concept of Kullback-Leibler divergence (also known as *relative entropy*). For two distributions $\pi_1(p)$ and $\pi_2(p)$ over $p \in \Gamma^P$, the Kullback-Leibler divergence $\mathcal{D}(\pi_1||\pi_2)$ of $\pi_1(p)$ relative to $\pi_2(p)$ is defined by

$$\mathcal{D}(\pi_1||\pi_2) = \sum_{p \in \Gamma^P} \pi_1(p) \ln\left(\frac{\pi_1(p)}{\pi_2(p)}\right). \tag{26}$$

Our cost function for precision can be defined as follows.

Assumption 1. The time cost of choosing a distribution $\pi(p)$, for $p \in \Gamma^P$, is $\kappa\mathcal{D}(\pi||u)$, where u represents the uniform distribution $u(p) = \frac{1}{\#P}$ for $p \in \Gamma^P$.

Here κ represents the marginal cost of entropy reduction, in units of labor time. The cost function in Assumption 1 can also be written as follows:

$$\kappa\mathcal{D}(\pi||u) = \kappa\left(\ln(\#P) + \sum_{j=1}^{\#P} \pi^j \ln(\pi^j)\right) \tag{27}$$

This cost function is nonnegative and convex.[10] It takes its maximum value, $\kappa \ln(\#P) > 0$, for any distribution that places all probability on a single price $p \in \Gamma^P$. It takes its minimum value, zero, for a uniform distribution.[11] Thus Assumption 1 means that decision costs are maximized by perfect precision and minimized by perfect randomness.

[8] This term comes from game theory; see Van Damme (1991), Chapter 4.

[9] See also Marsili 1999, Baron *et al.* 2002, and Matejka and McKay 2011.

[10] Cover and Thomas (2006), Theorem 2.7.2.

[11] If π is uniform, then $\pi(p) = 1/\#P$ for all $p \in \Gamma^P$, which implies $\sum_{j\in\Gamma^P} \pi(p)\ln(\pi(p)) = -\ln(\#P)$.

This cost function implies that the price choice is distributed as a multinomial logit. Suppose a firm at time t has already decided to update its price, and is now considering which new price P^j to choose from the finite grid $\Gamma^P \equiv \left\{P^1, P^2, ...P^{\#P}\right\}$. It will optimally choose a price distribution that maximizes firm value, net of computational costs (which we convert to nominal terms by multiplying by the wage):

$$\tilde{V}_t(A) = \max_{\pi^j} \sum_{j=1}^{\#P} \pi^j V_t(P^j, A) - \kappa W_t \left(\ln(\#P) + \sum_{j=1}^{\#P} \pi^j \ln(\pi^j) \right) \quad \text{s.t.} \sum_{j=1}^{\#P} \pi^j = 1 \tag{28}$$

The first-order condition for π^j is

$$V^j - \kappa W_t(1 + \ln \pi^j) - \mu = 0,$$

where μ is the multiplier on the constraint. Some rearrangement yields:

$$\pi^j = \exp\left(\frac{V^j}{\kappa W_t} - 1 - \frac{\mu}{\kappa W_t}\right). \tag{29}$$

Since the probabilities sum to one, we have $\exp\left(1 + \frac{\mu}{\kappa W_t}\right) = \sum_j \exp\left(\frac{V^j}{\kappa W_t}\right)$. Therefore the optimal probabilities (29) reduce to the logit formula (10).

By calculating the logarithm of π^j from (29), and plugging it into the objective, we can also obtain a simple analytical formula for the value function:

$$\tilde{V}_t(A) = \kappa W_t \ln\left(\frac{1}{\#P}\sum_{k=1}^{\#P} \exp\left(\frac{V_t(P^k, A)}{\kappa W_t}\right)\right). \tag{30}$$

This solution is convenient, since it means we can avoid doing numerical maximization in the step when we solve for the value of adjusting to a new price.

Thus, this version of our framework involves a cost of price adjustment, but implies that it should be interpreted as a cost of managerial effort rather than the more standard "menu cost" interpretation in terms of labor effort for the physical task of altering the posted price. Of course, if we choose to interpret the logit choice distribution as the result of costly managerial time, these costs should be subtracted out of the value function. In the description of the firm's problem, the expected value of adjustment, previously defined by (13), is now given by

$$D(P, A, \Omega) \equiv \tilde{V}(A, \Omega) - V(P, A, \Omega) \tag{31}$$

The managerial costs of adjustment are netted out of $\tilde{V}$, as we see in problem (28).

3.3 Choosing the timing of adjustment

By defining costs in terms of the Kullback-Leibler divergence of the price distribution, relative to a uniform distribution, we are penalizing any variation in the probability of one price relative to another. Next, we set up an analogous cost function that *penalizes variation in the probability of adjusting at any given time, relative to another.* Since the time to next adjustment could be arbitrarily far in the future, it makes no sense to penalize variation in the probability of actual arrival times relative to a uniform distribution (which would have unbounded support, implying an improper distribution). Instead, it is natural to penalize variation in the *arrival rate* of the adjustment time– in other words, to compare the adjustment time to a Poisson process.

Now, suppose the time period is sufficiently short so that we can approximately ignore multiple adjustments within a single period. If the firm adjusts its price at time t, it obtains the value gain $D_t(P_{it}, A_{it})$ defined in (31). Suppose it adjusts its price with probability λ_t. We measure the cost of this adjustment probability in terms of Kullback-Leibler divergence, relative to some arbitrary Poisson process with arrival rate $\bar{\lambda}$. In other words, we make the following assumption:

Assumption 2. Choosing to adjust with probability $\lambda_t \in [0, 1]$ in period t incurs the following time cost in period t:

$$\kappa\mathcal{D}((\lambda_t, 1 - \lambda_t)||(\bar{\lambda}, 1 - \bar{\lambda}))$$

for some constant $\bar{\lambda} \in [0, 1]$.

Here again, κ is the marginal cost of entropy reduction. Since the decision to adjust or not in any given period is a binary decision, Assumption 2 states that the decision cost in that period depends on the relative entropy of a binary decision with probabilities $(\lambda_t, 1 - \lambda_t)$, relative to another binary decision with probabilities $(\bar{\lambda}, 1 - \bar{\lambda})$.

In other words, what we are doing here is to benchmark the state-dependent price adjustment process $\lambda(L)$ in terms of the state-independent Calvo framework. This is a natural way to penalize variability in the distribution of a random time, just as comparing to a uniform distribution penalizes variability in the distribution of possible prices. Since a Calvo model can be defined at any arbitrary adjustment rate $\bar{\lambda}$, this setup implies the existence of one free parameter that measures the speed of decision making, in addition to the parameter κ^{-1} that measures the accuracy of decision making.

Given this cost function, the optimal adjustment probability satisfies

$$G_t(P_{it}, A_{it}) \;=\; \max_{\lambda_t} D_t(P_{it}, A_{it})\lambda_t - \kappa W_t \left[\lambda_t \ln\left(\frac{\lambda_t}{\bar{\lambda}}\right) + (1 - \lambda_t)\ln\left(\frac{1 - \lambda_t}{1 - \bar{\lambda}}\right)\right] \tag{32}$$

The first order condition is

$$D_t(P_{it}, A_{it}) = \kappa W_t \left[\ln \lambda_t + 1 - \ln \bar{\lambda}_t - \ln(1 - \lambda_t) - 1 + \ln(1 - \bar{\lambda})\right] \tag{33}$$

which simplifies to

$$\frac{\lambda_t}{1 - \lambda_t} = \frac{\bar{\lambda}}{1 - \bar{\lambda}} \exp\left(\frac{D_t}{\kappa W_t}\right) \tag{34}$$

Note that in continuous time, $\frac{\lambda_t}{1-\lambda_t} \to \lambda_t$, so (34) implies a well-defined continuous-time limit:

$$\lambda_t = \bar{\lambda} \exp\left(\frac{D_t}{\kappa W_t}\right) \in [0, \infty). \tag{35}$$

Alternatively, for a non-negligible discrete time interval, we can solve (34) to obtain

$$\lambda_t \equiv \lambda\left(\frac{D_t}{\kappa W_t}\right) \;=\; \frac{\bar{\lambda}}{\bar{\lambda} + (1 - \bar{\lambda})\exp\left(\frac{-D_t}{\kappa W_t}\right)} \tag{36}$$

$$= \; \frac{\bar{\lambda}\exp\left(\frac{E^\pi V_t}{\kappa W_t} - \mathcal{D}(\pi||u)\right)}{\bar{\lambda}\exp\left(\frac{E^\pi V_t}{\kappa W_t} - \mathcal{D}(\pi||u)\right) + (1 - \bar{\lambda})\exp\left(\frac{V_{it}}{\kappa W_t}\right)} \in [0, 1]. \tag{37}$$

This is the same weighted binary logit we obtained in Section 2.2.2, except that we are now explicitly subtracting off the costs of choosing the optimal price if the firm chooses to adjust.

Note also that this is the same weighted binary logit derived by Woodford (2009).[12] The free parameter $\bar{\lambda}$ measures the rate of decision making; concretely, the probability of adjustment in one discrete time period is $\bar{\lambda}$ when the firm is indifferent between adjusting and not adjusting, that is, when $D_t = 0$.

The value function $G_t(P, A)$ represents the expected gains from adjustment, net of the adjustment costs. Here again, we can explicitly solve for the value function. Rearranging the first-order conditions above, we have

$$\frac{1-\lambda_t}{1-\bar{\lambda}} = \frac{\lambda_t}{\bar{\lambda}} \exp\left(\frac{-D_t}{\kappa W_t}\right) = \left(1 - \bar{\lambda} + \bar{\lambda} \exp\left(\frac{D_t}{\kappa W_t}\right)\right)^{-1} \tag{38}$$

Plugging these formulas into the objective function, the value of problem (32) is

$$G_t(P, A) = \kappa W_t \ln\left(1 - \bar{\lambda} + \bar{\lambda} \exp\left(\frac{D_t(P, A)}{\kappa W_t}\right)\right). \tag{39}$$

Again, this solution conveniently allows us to avoid a numerical maximization step when we solve the firm's problem.

3.4 Recursive formulation of the firm's problem

Given these results on optimal decision-making under control costs, the firm's problem can be written in a fully recursive form, as follows.

$$V(P, A, \Omega) \quad = \quad U(P, A, \Omega) + \beta E\left\{ \tfrac{P(\Omega')C(\Omega')^{-\gamma}}{P(\Omega)C(\Omega)^{-\gamma}} \left[V(P, A', \Omega') + G(P, A', \Omega')\right] \Big| A, \Omega \right\}, \tag{40}$$

where

$$\begin{aligned} G(P, A', \Omega') \quad &\equiv \quad \max_{\lambda} \lambda D(P, A', \Omega') - W(\Omega')\kappa \mathcal{D}\left((\lambda, 1-\lambda) \,||\, (\bar{\lambda}, 1-\bar{\lambda})\right) \\ &= \quad \kappa W(\Omega') \ln\left(1 - \bar{\lambda} + \bar{\lambda} \exp\left(\frac{D(P, A', \Omega')}{\kappa W(\Omega')}\right)\right), \end{aligned} \tag{41}$$

$$D(P, A', \Omega') \quad \equiv \quad \tilde{V}(A', \Omega') - V(P, A', \Omega'), \tag{42}$$

and

$$\begin{aligned} \tilde{V}(A', \Omega') \quad &\equiv \quad \max_{\pi^j} \sum_{j=1}^{\#P} \pi^j V(P^j, A', \Omega') - W(\Omega')\kappa \left(\ln(\#P) + \sum_{j=1}^{\#P} \pi^j \ln(\pi^j) \right) \\ &= \quad \kappa W(\Omega') \ln\left(\frac{1}{\#P} \sum_{j=1}^{\#P} \exp\left(\frac{V(P^j, A', \Omega')}{\kappa W(\Omega')} \right) \right). \end{aligned} \tag{43}$$

[12] The published version of Woodford's paper only states a first-order condition equivalent to (34). Earlier working paper versions pointed out that the first-order condition implies a logit hazard of the form (36).

The terms inside the expectation in the Bellman equation represent the value V of continuing without adjustment, plus the flow of expected gains G due to adjustment. Note that the function G is *known analytically* in terms of the function $D = \tilde{V} - V$. But likewise, $\tilde{V}$ is *known analytically* in terms of the function V. In other words, running numerical backwards induction in this context is especially simple, because all the maximization steps can be performed analytically.

The price process associated with (40) is

$$P_{it} = \begin{cases} P^j \in \Gamma^P & \text{with probability} \quad \lambda\left(\frac{D(\widetilde{P}_{it}, A_{it}, \Omega_t)}{\kappa W(\Omega_t)}\right) \pi(P^j | A_{it}, \Omega_t) \\ \widetilde{P}_{it} \equiv P_{i,t-1} & \text{with probability} \quad 1 - \lambda\left(\frac{D(\widetilde{P}_{it}, A_{it}, \Omega_t)}{\kappa W(\Omega_t)}\right) \end{cases} . \tag{44}$$

Here, the adjustment probability λ is given by (36), and the price distribution π is given by (10). Equation (44) is written with time subscripts for additional clarity.

4 Computation

4.1 Outline of algorithm

Computing this model is challenging due to heterogeneity: at any time t, firms will face different idiosyncratic shocks A_{it} and will be stuck at different prices P_{it}. The reason for the popularity of the Calvo model is that even though firms have many different prices, up to a first-order approximation only the average price matters for equilibrium. Unfortunately, this property does not hold in general, and in the current context, we need to treat all equilibrium quantities explicitly as functions of the distribution of prices and productivity across the economy, and we must calculate the dynamics of this distribution over time.

We address this problem by implementing Reiter's (2009) solution method for dynamic general equilibrium models with heterogeneous agents and aggregate shocks. As a first step, Reiter's algorithm calculates the steady state general equilibrium that obtains in the absence of aggregate shocks. Idiosyncratic shocks are still active, but are assumed to have converged to their ergodic distribution, so an aggregate steady state means that $z = 0$, and Ψ, π, C, R, N, and w are all constant. To solve for this steady state, we will assume that real prices and productivities always lie on a fixed grid $\Gamma \equiv \Gamma^P \times \Gamma^a$, where $\Gamma^p \equiv \{p^1, p^2, ...p^{\#^p}\}$ and $\Gamma^a \equiv \{a^1, a^2, ...a^{\#^a}\}$ are logarithmically-spaced grids of possible values of p_{it} and A_{it}, respectively. We can then think of the steady state value function as a matrix $\mathbf{V}$ of size $\#^p \times \#^a$ comprising the values $v^{jk} \equiv v(p^j, a^k)$ associated with the prices and productivities $\left(p^j, a^k\right) \in \Gamma$. Likewise, the price distribution can be viewed as a $\#^p \times \#^a$ matrix $\mathbf{\Psi}$ in which the row j, column k element Ψ^{jk} represents the fraction of firms in state (p^j, a^k) at the time of production. Given this discretized representation, we can calculate steady state general equilibrium by guessing the aggregate wage level, then solving the firm's problem by backwards induction on the grid Γ, then updating the conjectured wage, and iterating to convergence.

In a second step, Reiter's method constructs a linear approximation to the dynamics of the discretized model, by perturbing it around the steady state general equilibrium on a point-by-point basis. The method recognizes that the Bellman equation and the distributional dynamics can be interpreted as a large system of nonlinear first-order autonomous difference equations that define the aggregate dynamics. For example, away from steady state, the Bellman equation relates the $\#^p \times \#^a$ matrices $\mathbf{V}_t$ and $\mathbf{V}_{t+1}$ that represent the value function at times t and $t+1$. The row j, column k element of $\mathbf{V}_t$ is $v_t^{jk} \equiv v_t(p^j, a^k) \equiv v(p^j, a^k, \Xi_t)$, for $\left(p^j, a^k\right) \in \Gamma$. Given this representation, we no longer need to think of the Bellman equation as a functional equation that defines $v(p, a, \Xi)$ for all possible idiosyncratic and aggregate states p, a, and Ξ; instead, we simply treat it as a system of $\#^p\#^a$ expectational difference equations that determine the dynamics of the $\#^p\#^a$ variables v_t^{jk}. We linearize this large system of difference equations numerically, and then solve for the saddle-path stable solution of our linearized model using the QZ decomposition, following Klein (2000).

The beauty of Reiter's method is that it combines linearity and nonlinearity in a way appropriate for the model at hand. In the context of price setting, aggregate shocks are likely to be less relevant for individual firms' decisions than idiosyncratic shocks; Klenow and Kryvstov (2008), Golosov and Lucas (2007), and Midrigan (2008) all argue that firms' prices are driven primarily by idiosyncratic shocks. To deal with these big firm-specific shocks, we treat functions of idiosyncratic states in a fully nonlinear way, by calculating them on a grid. But this grid-based solution can also be regarded as a large system of nonlinear equations, with equations specific to each of the grid points. When we linearize each of these equations with respect to the aggregate dynamics, we recognize that aggregate changes are unlikely to affect individual value functions in a strongly nonlinear way. That is, we are implicitly assuming that aggregate shocks z_t and changes in the distribution $\mathbf{\Psi}_t$ have sufficiently smooth impacts on individual values that a linear treatment of these effects suffices. On the other hand, we need *not* start from any assumption of approximate aggregation like that required for the Krusell and Smith (1998) method, nor do we need to impose any particular functional form on the distribution $\mathbf{\Psi}$.

Describing the distributional dynamics involves defining various matrices related to quantities on the grid Γ. From here on, we use bold face to identify matrices, and superscripts to identify notation related to grids. Matrices associated with the grid Γ are defined so that row j relates to the price $p^j \in \Gamma^p$, and column k relates to the productivity $a^k \in \Gamma^a$. Besides the value function matrix $\mathbf{V}_t$, we also define matrices $\mathbf{D}_t$, $\mathbf{G}_t$, and $\mathbf{\Lambda}_t$, to represent the functions d_t, g_t, and $\lambda(d_t/w_t)$ at points on the grid Γ. The distribution at the time of production is given by $\mathbf{\Psi}_t$, with elements Ψ_t^{jk} representing the fraction of firms with real price $p_{it} \equiv P_{it}/P_t = p^j$ and productivity $A_{it} = a^k$ at the time of production. We also define the beginning of period distribution $\widetilde{\mathbf{\Psi}}_t$, with elements $\widetilde{\Psi}_t^{jk}$ representing the fraction of firms with real price $\widetilde{p}_{it} \equiv \widetilde{P}_{it}/P_t = p^j$ and productivity $A_{it} = a^k$ at the beginning of the period. Shortly we will define the transition matrices that govern the relationships between all these objects.

4.2 The discretized model

In the discretized model, the value function $\mathbf{V}_t$ is a matrix of size $\#^p \times \#^a$ with elements $v_t^{jk} \equiv v_t(p^j, a^k) \equiv v(p^j, a^k, \Xi_t)$ where $(p^j, a^k) \in \Gamma$. Other relevant $\#^p \times \#^a$ matrices include the adjustment values $\mathbf{D}_t$, the probabilities $\mathbf{\Lambda}_t$, and the expected gains $\mathbf{G}_t$, with (j, k) elements given by[13]

$$d_t^{jk} \equiv d_t(p^j, a^k) \equiv E_t^\pi v_t(p, a^k) - v_t(p^j, a^k) \tag{45}$$

$$\lambda_t^{jk} \equiv \lambda\left(d_t^{jk}/w_t\right) \tag{46}$$

[13]Actually, (46) is a simplified description of λ_t^{jk}. While (46) implies that λ_t^{jk} represents the function $\lambda(L)$ evaluated at the log price grid point p^j and log productivity grid pointa^k, in our computations λ_t^{jk} actually represents the *average* of $\lambda(L)$ over all log prices in the interval $\left(\frac{p^{j-1}+p^j}{2}, \frac{p^j+p^{j+1}}{2}\right)$, given log productivity a^k. Calculating this average requires interpolating the function $d_t(p, a^k)$ between price grid points. Defining λ_t^{jk} this way ensures differentiability with respect to changes in the aggregate state Ω_t.

$$g_t^{jk} \equiv \lambda_t^{jk} d_t^{jk} \tag{47}$$

Finally, we also define a matrix of logit probabilities $\mathbf{\Pi}_t$, which has its (j,k) element given by

$$\pi_t^{jk} = \pi_t(p^j|a^k) \equiv \frac{\exp\left(v_t^{jk}/(\kappa w_t)\right)}{\sum_{n=1}^{\#p} \exp\left(v_t^{jn}/(\kappa w_t)\right)}$$

which is the probability of choosing real price p^j conditional on productivity a^k if the firm decides to adjust its price at time t.

The control cost version of the model differs only in the definitions of d_t^{jk} and g_t^{jk}. Equations (45) and (47) are replaced by

$$d_t^{jk} \equiv d_t(p^j, a^k) \equiv E_t^{\pi} v_t(p, a^k) - v_t(p^j, a^k) - w_t \kappa \mathcal{D}_{\pi} \tag{48}$$

$$g_t^{jk} \equiv \lambda_t^{jk} d_t^{jk} - w_t \kappa \mathcal{D}_{\lambda} \tag{49}$$

We can now write the discrete Bellman equation and the discrete distributional dynamics in a precise way. First, consider how the beginning-of-period distribution $\widetilde{\mathbf{\Psi}}_t$ is derived from the lagged distribution $\mathbf{\Psi}_{t-1}$. Idiosyncratic productivities A_i are driven by an exogenous Markov process, which can be defined in terms of a matrix $\mathbf{S}$ of size $\#^a \times \#^a$. The row m, column k element of $\mathbf{S}$ represents the probability

$$S^{mk} = prob(A_{it} = a^m | A_{i,t-1} = a^k)$$

Also, beginning-of-period real prices are, by definition, adjusted for inflation. Ignoring grids, the time $t-1$ real price $p_{i,t-1}$ would deflated to $\widetilde{p}_{it} \equiv p_{i,t-1}/\pi_t \equiv p_{i,t-1}P_{t-1}/P_t$ at the beginning of t. To keep prices on the grid, we define a $\#^p \times \#^p$ Markov matrix $\mathbf{R}_t$ in which the row m, column l element is

$$R_t^{ml} \equiv prob(\widetilde{p}_{it} = p^m | p_{i,t-1} = p^l)$$

When the deflated price $p_{i,t-1}/\pi_t$ falls between two grid points, matrix $\mathbf{R}_t$ must round up or down stochastically. Also, if $p_{i,t-1}/\pi_t$ lies outside the smallest and largest element of the grid, then $\mathbf{R}_t$ must round up or down to keep prices on the grid.[14] Therefore we construct $\mathbf{R}_t$ according to

$$R_t^{ml} = prob(\widetilde{p}_{it} = p^m | p_{i,t-1} = p^l, \pi_t) = \begin{cases} 1 & \text{if } \pi_t^{-1} p^l \le p^1 = p^m \\ \frac{\pi_t^{-1} p^l - p^{n-1}}{p^n - p^{n-1}} & \text{if } p^1 < p^m = \min\{p \in \Gamma^p : p \ge \pi_t^{-1} p^l\} \\ \frac{p^{n+1} - \pi_t^{-1} p^l}{p^{n+1} - p^n} & \text{if } p^1 \le p^m = \max\{p \in \Gamma^p : p < \pi_t^{-1} p^l\} \\ 1 & \text{if } \pi_t^{-1} p^l > p^{\#^p} = p^m \\ 0 & \text{otherwise} \end{cases} \tag{50}$$

[14]In other words, we assume that any nominal price that would have a real value less than p^1 after inflation is automatically adjusted upwards so that its real value is p^1. This assumption is made for numerical purposes only, and has a negligible impact on the equilibrium as long as we choose a sufficiently wide grid Γ^p. If we were to compute examples with trend deflation, we would need to make an analogous adjustment to prevent real prices from exceding the maximum grid point $p^{\#^p}$.

Combining the adjustments of prices and productivities, we can calculate the beginning-of-period distribution $\widetilde{\boldsymbol{\Psi}}_t$ as a function of the lagged distribution of production prices $\boldsymbol{\Psi}_{t-1}$:

$$\widetilde{\boldsymbol{\Psi}}_t = \mathbf{R}_t * \boldsymbol{\Psi}_{t-1} * \mathbf{S}'$$

where $*$ represents ordinary matrix multiplication. The simplicity of this equation comes partly from the fact that the exogenous shocks to A_{it} are independent of the inflation adjustment that links $\widetilde{p}_{it}$ with p_{it-1}. Also, exogenous shocks are represented from left to right in the matrix $\boldsymbol{\Psi}_t$, so that their transitions can be treated by right multiplication, while policies are represented vertically, so that transitions related to policies can be treated by left multiplication.

To calculate the effects of price adjustment on the distribution, let $\mathcal{E}_{pp}$ and $\mathcal{E}_{pa}$ by matrices of ones of size $\#P \times \#P$ and $\#P \times \#a$, respectively. Now suppose a firm has beginning-of-t price $\widetilde{p}_{it} \equiv \widetilde{P}_{it}/P_t = p^j \in \Gamma^p$ and productivity $A_{it} = a^k \in \Gamma^a$. This firm will adjust its production price with probability λ_t^{jk}, or will leave it unchanged ($p_{it} = \widetilde{p}_{it} = p^j$) with probability $1 - \lambda_t^{jk}$. If adjustment occurs, the probabilities of choosing all possible prices are given by the matrix $\boldsymbol{\Pi}_t$. Therefore we can calculate distribution $\boldsymbol{\Psi}_t$ from $\widetilde{\boldsymbol{\Psi}}_t$ as follows:

$$\boldsymbol{\Psi}_t = (\mathbf{E}_{pa} - \boldsymbol{\Lambda}) \;.*\; \widetilde{\boldsymbol{\Psi}}_t + \boldsymbol{\Pi}_t \;.*\; (\mathbf{E}_{pp} * (\boldsymbol{\Lambda} \;.*\; \widetilde{\boldsymbol{\Psi}}_t)) \tag{51}$$

where (as in MATLAB) the operator $.*$ represents element-by-element multiplication, and $*$ represents ordinary matrix multiplication.

The same transition matrices $\mathbf{R}$ and $\mathbf{S}$ show up when we write the Bellman equation in matrix form. Let $\mathbf{U}_t$ be the $\#^p \times \#^a$ matrix of current payoffs, with elements

$$u_t^{jk} \equiv \left(p^j - \frac{w_t}{a^k}\right)(p^j)^{-\epsilon} C_t \tag{52}$$

for $\left(p^j, a^k\right) \in \Gamma$. Then the Bellman equation is
Dynamic general equilibrium Bellman equation, matrix version:

$$\mathbf{V}_t = \mathbf{U}_t + \beta E_t \left\{ \frac{u'(C_{t+1})}{u'(C_t)} \left[\mathbf{R}'_{t+1} * (\mathbf{V}_{t+1} + \mathbf{G}_{t+1}) * \mathbf{S}\right] \right\} \tag{53}$$

The expectation E_t in the Bellman equation refers only to the effects of the time $t+1$ aggregate shock z_{t+1}, because the shocks and dynamics of the idiosyncratic state $(p^j, a^k) \in \Gamma$ are completely described by the matrices $\mathbf{R}'_{t+1}$ and $\mathbf{S}$. Note that since the Bellman equation iterates backwards in time, its transitions are represented by $\mathbf{R}'$ and $\mathbf{S}$, whereas the distributional dynamics iterate forward in time and therefore contain $\mathbf{R}$ and $\mathbf{S}'$.

While equilibrium seems to involve a very complex system of equations, the steady state is easy to solve because it reduces to a small scalar fixed-point problem, which is the first step of Reiter's (2009) method. This first step is discussed in the next subsection. The second step of the method, in which we linearize all equilibrium equations, is discussed in subsection 3.4.

4.3 Step 1: steady state

In the aggregate steady state, the shocks are zero, and the distribution takes some unchanging value $\boldsymbol{\Psi}$, so the state of the economy is constant: $\Xi_t \equiv (z_t, \Psi_{t-1}) = (0, \boldsymbol{\Psi}) \equiv \Xi$. We indicate the steady state of all equilibrium objects by dropping the time subscript t, so the steady state value function $\mathbf{V}$ has elements $v^{jk} \equiv v(p^j, a^k, \Xi) \equiv v(p^j, a^k)$.

Long run monetary neutrality in steady state implies that the rate of nominal money growth equals the rate of inflation:

$$\mu = \pi$$

Morever, the Euler equation reduces to

$$\pi = \beta R$$

Since the interest rate and inflation rate are observable, together they determine the required parameterization of β. The steady-state transition matrix $\mathbf{R}$ is known, since it depends only on steady state inflation π.

We can then calculate general equilibrium as a one-dimensional root-finding problem: guessing the wage w, we have enough information to solve the Bellman equation and the distributional dynamics.[15] Knowing the steady state aggregate distribution, we can construct the real price level, which must be one. Thus finding a value of w at which the real price level is one amounts to finding a steady state general equilibrium.

More precisely, for any w, we can calculate

$$C = \left(\frac{\chi}{w}\right)^{1/\gamma} \tag{54}$$

and then construct the matrix $\mathbf{U}$ with elements

$$u^{jk} \equiv \left(p^j - \frac{w}{a^k}\right)(p^j)^{-\epsilon} C \tag{55}$$

We then find the fixed point of the value function:

$$\mathbf{V} = \mathbf{U} + \beta \mathbf{R}' * (\mathbf{V} + \mathbf{G}) * \mathbf{S} \tag{56}$$

together with the logit probability function $\mathbf{\Pi}$, with elements

$$\pi^{jk} = \pi_t(p^j|a^k) \equiv \frac{\exp\left(v^{jk}/(\kappa w)\right)}{\sum_{n=1}^{\#p} \exp\left(v^{jn}/(\kappa w)\right)}$$

We can then find the steady state distribution as the fixed point of

$$\mathbf{\Psi} = (\mathbf{E}_{pa} - \mathbf{\Lambda}) \;.*\; \widetilde{\mathbf{\Psi}} + \mathbf{\Pi} \;.*\; (\mathbf{E}_{pp} * (\mathbf{\Lambda} \;.*\; \widetilde{\mathbf{\Psi}})) \tag{57}$$

$$\widetilde{\mathbf{\Psi}} = \mathbf{R} * \mathbf{\Psi} * \mathbf{S}' \tag{58}$$

Finally, we check whether

$$1 \;=\; \sum_{j=1}^{\#^p} \sum_{k=1}^{\#^a} \Psi^{jk} \left(p^j\right)^{1-\epsilon} \equiv p(w) \tag{59}$$

If so, an equilibrium value of w has been found.

[15] There are other, equivalent ways of describing the root-finding problem: for example, we could begin by guessing C. Guessing w is convenient since we know that in a representative-agent, flexible-price model, we have $w = \frac{\epsilon-1}{\epsilon}$. This suggests a good starting value for the heterogeneous-agent, sticky-price calculation.

4.4 Step 2: linearized dynamics

Given the steady state, the general equilibrium dynamics can be calculated by linearization. To do so, we eliminate as many variables from the equation system as we can. For additional simplicity, we assume linear labor disutility, $x(N) = \chi N$. Thus the first-order condition for labor reduces to $\chi = w_t u'(C_t)$, so we don't actually need to solve for N_t in order to calculate the rest of the equilibrium.[16] We can then summarize the general equilibrium equation system in terms of the exogenous shock process z_t, the lagged distribution of idiosyncratic states $\mathbf{\Psi}_{t-1}$, which is the endogenous component of the time t aggregate state; and finally the endogenous 'jump' variables including $\mathbf{V}_t$, $\mathbf{\Pi}_t$, C_t, R_t, and π_t. The equation systems reduces to

$$z_t = \phi_z z_{t-1} + \epsilon_t^z \tag{60}$$

$$\mathbf{\Psi}_t = (\mathbf{E}_{pa} - \mathbf{\Lambda}_t)\ .*\ \widetilde{\Psi}_t + \mathbf{\Pi}_t\ .*\ (\mathbf{E}_{pp} * (\mathbf{\Lambda}_t\ .*\ \widetilde{\mathbf{\Psi}}_t)) \tag{61}$$

$$\mathbf{V}_t = \mathbf{U}_t + \beta E_t \left\{ \frac{u'(C_{t+1})}{u'(C_t)} \left[\mathbf{R}'_{t+1} * (\mathbf{V}_{t+1} + \mathbf{G}_{t+1}) * \mathbf{S} \right] \right\} \tag{62}$$

$$R_t^{-1} = \beta E_t \left(\frac{u'(C_{t+1})}{\pi_{t+1} u'(C_t)} \right) \tag{63}$$

$$1 = \sum_{j=1}^{\#^p} \sum_{k=1}^{\#^a} \Psi_t^{jk} \left(p^j\right)^{1-\epsilon} \tag{64}$$

If we now collapse all the endogenous variables into a single vector

$$\overrightarrow{X}_t \equiv \left(vec\left(\mathbf{\Psi}_{t-1}\right)', \quad vec\left(\mathbf{V}_t\right)', \quad C_t, \quad R_t, \quad \pi_t \right)'$$

then the whole set of expectational difference equations (60)-(64) governing the dynamic equilibrium becomes a first-order system of the following form:

$$E_t \mathcal{F}\left(\overrightarrow{X}_{t+1}, \overrightarrow{X}_t, z_{t+1}, z_t \right) = 0 \tag{65}$$

where E_t is an expectation conditional on z_t and all previous shocks.

To see that the variables in vector $\overrightarrow{X}_t$ are in fact the only variables we need, note that given π_t and π_{t+1}we can construct $\mathbf{R}_t$ and $\mathbf{R}_{t+1}$. Given $\mathbf{R}_t$, we can construct $\widetilde{\mathbf{\Psi}}_t = \mathbf{R}_t * \mathbf{\Psi}_{t-1} * \mathbf{S}'$ from $\mathbf{\Psi}_{t-1}$. Under linear labor disutility, we can calculate $w_t = \chi / u'(C_t)$, which gives us all the information needed to construct $\mathbf{U}_t$, with (j,k) element equal to $u_t^{jk} \equiv \left(p^j - \frac{w_t}{a^k}\right)(p^j)^{-\epsilon} C_t$. Finally, given $\mathbf{V}_t$ and $\mathbf{V}_{t+1}$ we can construct $\mathbf{\Pi}_t$, $\mathbf{D}_t$, and $\mathbf{D}_{t+1}$, and thus $\mathbf{\Lambda}_t$ and $\mathbf{G}_{t+1}$. Therefore the variables in $\overrightarrow{X}_t$ and z_t are indeed sufficient to evaluate the system (60)-(64).

Finally, if we linearize system $\mathcal{F}$ numerically with respect to all its arguments to construct the Jacobian matrices $\mathcal{A} \equiv D_{\overrightarrow{X}_{t+1}} \mathcal{F}$, $\mathcal{B} \equiv D_{\overrightarrow{X}_t} \mathcal{F}$, $\mathcal{C} \equiv D_{z_{t+1}} \mathcal{F}$, and $\mathcal{D} \equiv D_{z_t} \mathcal{F}$, then we obtain the following first-order linear expectational difference equation system:

$$E_t \mathcal{A} \Delta \overrightarrow{X}_{t+1} + \mathcal{B} \Delta \overrightarrow{X}_t + E_t \mathcal{C} z_{t+1} + \mathcal{D} z_t = 0 \tag{66}$$

where Δ represents a deviation from steady state. This system has the form considered by Klein (2000), so we solve our model using his QZ decomposition method.[17]

[16] The assumption $x(N) = \xi N$ is not essential; the more general case with nonlinear labor disutility simply requires us to simulate a larger equation system that includes N_t.

[17] Alternatively, the equation system can be rewritten in the form of Sims (2001). We chose to implement the Klein method because it is especially simple and transparent to program.

5 Results

We next describe the calibration of the model and report simulation results. We describe the model's steady state implications for microdata on price adjustments, both at a low inflation rate, and as the rate of trend inflation is substantially increased. We also analyze the macroeconomic implications for the effects of monetary policy shocks. Our focus throughout is on understanding the behavior of our model of error-prone price setting. Therefore our simulations compare our main model, in which both the timing of price adjustment and the size of the adjustments are subject to errors, with alternative models that shut down one or the other margin of errors.[18] The simulation is performed at monthly frequency, and all data and model statistics are monthly unless stated otherwise.

5.1 Parameters

The key parameters we must estimate to understand the behavior of our model are the rate and noise parameters of the decision process, $\bar{\lambda}$ and κ. We estimate these two parameters to match two steady-state properties of the price process: the average rate of adjustment, and the histogram of nonzero log price adjustments, in the Dominick's supermarket dataset described in Midrigan (2011).[19] More precisely, let h be a vector of length $\#h$ representing the frequencies of nonzero log price adjustments in a histogram with $\#h$ fixed bins.[20] We choose $\bar{\lambda}$ and κ to minimize the following distance criterion:

$$\text{distance} \;=\; \sqrt{\#h}\,||\lambda_{model} - \lambda_{data}|| + ||h_{model} - h_{data}|| \tag{67}$$

where $||\bullet||$ represents the Euclidean norm, λ_{model} and λ_{data} represent the average frequency of price adjustment in the simulated model and in the Dominick's dataset, and h_{model} and h_{data} are the vectors of bin frequencies for nonzero price adjustments in the model and the data.[21] Clearly these features of the data are informative about the two parameters, since $\bar{\lambda}$ will shift the frequency of adjustment and κ will spread the distribution of price adjustments.

The rest of the parameterization is less crucial for our purposes. Therefore, for comparability, we take our utility parameterization directly from Golosov and Lucas (2007). Thus, we set the discount factor to $\beta = 1.04^{-1/12}$. Consumption utility is CRRA, $u(C) = \frac{1}{1-\gamma}C^{1-\gamma}$, with $\gamma = 2$. Labor disutility is linear, $x(N) = \chi N$, with $\chi = 6$. The elasticity of substitution in the consumption aggregator is $\epsilon = 7$. Finally, the utility of real money holdings is logarithmic, $v(m) = \nu \log(m)$, with $\nu = 1$. We assume productivity is AR(1) in logs: $\log A_{it} = \rho \log A_{it-1} + \varepsilon_t^a$, where ε_t^a is a mean-zero, normal, *iid* shock. We take the autocorrelation parameter from Blundell and Bond (2000), who estimate it from a panel of 509 US manufacturing companies over 8 years, 1982-1989. Their preferred estimate is 0.565 on an annual basis, which implies ρ around 0.95 at monthly frequency. The variance of log productivity is $\sigma_a^2 = (1-\rho^2)\sigma_\varepsilon^2$, where σ_ε^2 is the variance

[18] Alternatively, we could compare our main model to more standard price adjustment models. However, in Costain and Nakov (2011C), we have already compared the model with errors in the size of price adjustment to the Calvo and menu cost models. We refer readers to that paper for comparable tables and graphs documenting the behavior of those models.

[19] The weekly adjustment rate in the Dominick's data is aggregated to a monthly rate for comparability with the model.

[20] See Figure **??**, which compares these histograms in the data and in all specifications of our model.

[21] Since the Euclidean norm of a vector scales with the square root of the number of elements, we scale the first term by $\#h$ to place roughly equal weight on the two components of the distance measure.

of the innovation ε_t^a. We set the standard deviation of log productivity to $\sigma_a = 0.06$, which is the standard deviation of "reference costs" estimated by Eichenbaum, Jaimovich, and Rebelo (2011). The rate of money growth is set to match the roughly 2% annual inflation rate observed in the Dominick's dataset.

In our results we will compare six specifications. The specifications with errors in the size of price adjustments, but not in their timing, are labelled "PPS", for "precautionary price stickiness". The specifications with errors in the timing of price adjustments, but not in their size, are labelled "Woodford", since the adjustment hazard takes the functional form derived in Woodford (2009). The specifications with both types of errors are labelled "nested". For all three cases, we report the model based on control costs, as well as a model that imposes errors of logit functional form exogenously, without deriving these errors from control costs.

Parameter estimates for these six specifications are reported in Table 1. Note that the PPS specification has only one free parameter: the level of noise κ_π in the pricing decision. The Woodford model has two free parameters: the rate parameter $\bar{\lambda}$, and the level of noise κ_λ in the timing decision. The nested model features the same two free parameters, except that the noise parameter now applies both to the timing and pricing decisions ($\kappa_\pi = \kappa_\lambda \equiv \kappa$).[22] Typically there is not much difference in the estimated parameters between the logit and control cost specifications, except for the "PPS" case, where the estimated noise is much smaller under control costs than it is under an exogenous logit. Overall the estimates imply a low level of noise, compared to values typically reported in experimental studies ($\kappa = 0$ would represent errorless choice). The rate parameter $\bar{lambda}$ is estimated to be lower than the observed adjustment frequency in the Woodford specification, but is twice as high as the observed adjustment frequency in the main model, marked "nested control". The combination of a high underlying adjustment rate, together with a low noise parameter, indicates a high degree of rationality in the estimate of the main model.

5.2 Results: microeconomics of price adjustment

The steady state behavior of the main model is illustrated in Fig. 1. The first panel of the figure illustrates the distribution of prices chosen conditional on productivity, $\pi(p|a)$; the axes show prices and costs (inverse productivity), expressed in log deviations from their unconditional means. As expected, the mean price chosen increases roughly one-for-one with cost, but the smooth bell-shape of the distribution conditional on a reflects the presence of errors. Similarly, the second panel shows the probability $\lambda(d(p,a)/(\kappa w))$ of price adjustment conditional on beginning-of-period price and productivity. Near the 45^o-line, the adjustment probability reaches a (strictly positive) minimum; moving away from the 45^o-line, it increases smoothly towards one. The third panel is a contour plot of the end-of-period distribution of prices and productivities, $\Psi(p,a)$. Dispersion in the horizontal direction represents variation in idiosyncratic productivity over time; dispersion in the vertical direction represents deviation from the conditionally-optimal price, caused either by failures to adjust in response to productivity shocks, or by errors when adjustment occurs. This distribution spreads out horizontally at the beginning of the period when new productivity shocks hit. The resulting distribution of adjusting firms is illustrated by the contour plot in the last panel of the figure. The most frequently observed adjustments

[22] It would also be interesting to allow the two noise parameters of the nested specification to differ, but we leave this for future work, since the simple cross-sectional statistics we are using may not suffice to identify these parameteres separately.

occur at firms whose prices deviate from their conditionally-optimal values by 5%-10%; firms with smaller deviations have little incentive to adjust, while firms with larger deviations are rare because adjustment usually occurs before a larger deviation is reached. The asymmetry observed in the density of adjustments reflects the fact that downward price errors (implying high sales at an unprofitably low price) are more costly than upward price errors.

This adjustment process is also illustrated in Fig. 2, which shows the distribution of losses $d(p, a)$ from nonadjustment, expressed as a percentage of average monthly revenue, at the beginning and end of the period. At the beginning of the period (shaded blue bars), the distribution of losses skews out to the right; losses of up to 8% of revenue are visible in the histogram. Adjustment eliminates some, but not all, of these largest losses. Thus, in the distribution of losses at the end of the period, some mass is shifted from the right tail of the distribution to the left end, where the loss is *negative*– that is, firms at the left end of this distribution are strictly better off not adjusting, because the risk of making a pricing error is too costly (this is the phenomenon we call "precautionary price stickiness"). Adjustment fails to completely eliminate the right tail of the distribution for two reasons: some firms that are expected to benefit from adjustment fail to adjust, and some firms that adjust make costly errors.

Table 2 and Fig. 3 also compare other specifications of the model. Table 2 reports statistics from the steady state of each specification of model, and the corresponding statistics from the Dominick's data. All specifications successfully match the 10.2% monthly adjustment frequency observed in the data. However, the nested specifications are much more successful in matching most of the statistics of the price distribution. The typical size of the adjustments (measured either as the mean absolute change, or the standard deviation of the adjustments) is too small in the Woodford model and in PPS-control, whereas it is too large in PPS-logit. The nested specification performs better on both these measures, and also in terms of the fraction of very small adjustments (that is, smaller than 5%). Likewise, the nested specification matches the high kurtosis of the data better than the other specifications do. The only reported statistic where the nested model performs less well is the fraction of positive adjustments, which is matched very well by the Woodford specification and by PPS-control.

These differences are clarified by graphing the histogram of nonzero log price adjustments, shown in Fig. 3 for the data (shaded blue bars) and for all six model specifications. The vector of bin frequencies for the 81 bars in these histograms is the object that enters the second term of the distance criterion used in the estimation. The distribution of nonzero adjustments in the data is characterized by a small peak of negative adjustments, a high peak of positive adjustments, and very fat tails. The implications of the PPS model (shown in the first panel of the figure) differ strongly between the exogenous logit and control cost specifications. As Table 1 showed, the estimated degree of noise is much lower when control costs are included. This makes sense: *ceteris paribus*, adjustment is less likely if adjusting implies a decision cost. Therefore, to match the same empirical frequency of adjustment in the logit and control cost specifications, price adjustment must be *less risky* (that is, subject to less noise) in the control cost case. The result is that our estimate of the control cost version of the PPS model has extremely low noise, resulting in behavior that is very close to full rationality. Therefore, the implied distribution of price adjustments resembles that in a fixed menu cost model, with two sharp spikes in the histogram representing increases or decreases occurring near a pair of (S,s) bands. On average, this implies much smaller price adjustments than those in the data, with little mass in the tails of the distribution.

In contrast, the exogenous logit version of the PPS model requires much more noise to produce the same average adjustment frequency. This implies a smoother, wider, more bell-shaped distribution than that observed in the data. In summary, the single free parameter of the PPS framework provides insufficient flexibility to match both the average frequency and the average size of price adjustments. In Costain and Nakov (2011C), for a different dataset with a zero average inflation rate, we reported an estimate of the PPS model that matched both the frequency and size of price adjustments well. But this finding was essentially coincidental; in the current dataset matching the mean adjustment frequency either implies price changes that are too small (assuming control costs) or too large (assuming an exogenous logit).

The Woodford specification, on the other hand, has two free parameters, so it may potentially work better in fitting the frequency and size of adjustments. However, with no errors in the chosen price, this specification implies a much tighter distribution of adjustments than those observed in the data: while the data show some price changes as large as $\pm 50\%$, our estimate of the Woodford specification implies no price changes larger than $\pm 20\%$. While a sufficiently high volatility of underlying costs would spread out the distribution of adjustments observed in this specification, by itself this would not necessarily reproduce the fat tails of the empirical distribution of price adjustments. In this sense, the nested specification fits the data better than the Woodford specification. Both models have just two parameters (since we are constraining the noise parameter in the timing decision to be the same as the noise parameter in the priceing decision). But the pricing errors present in the nested model make it easier to generate a wide, fat-tailed distribution than it is in the Woodford model. At the same time, the parameter $\bar{\lambda}$ helps ensure that the nested model gets the adjustment frequency right. Stated differently, the restriction $\kappa_\pi = 0$ imposed by the Woodford specification strongly constrains its ability to match the data, whereas the restriction $\kappa_\pi = \kappa_\lambda$ that we have maintained when estimating the nested specification does not appear to be very inconsistent with the data. While the main peak is smoother than that observed in the data, the nested model is quite successful both in reproducing the average size of price adjustments and in generating relatively fat tails.

Figures 4-6 show how the six specifications compare in reproducing several potentially puzzling aspects of the microdata. First, one might intuitively expect price adjustment hazards to increase with the time since last adjustment. But empirically, price adjustment hazards are *decreasing* with the time since adjustment, even after controlling for heterogeneity, as in Figure 4, where the blue shaded bars are the adjustment hazards reported by Nakamura and Steinsson (2008). Comparing the various versions of our model. we see that under Woodford's logit the adjustment hazard increases over time, since newly set prices are conditionally optimal, and subsequent inflation and productivity shocks gradually drive prices out of line with costs. In contrast, under the PPS-logit specification the adjustment hazard decreases very strongly with the time since last adjustment. This is a consequence of the relatively noisy decisions implied by the estimated parameters for this specification– prices adjust again quickly after a large error occurs. A similar effect exists in PPS-control and the nested models– the possibility of errors in price setting makes the adjustment hazard downward sloping. The downward slope is much milder than it was for PPS-logit, both because the noise in the pricing decision is smaller, and because errors in the *timing* of price adjustment imply that firms do not always respond immediately when they err in the *size* of their adjustments. Thus, PPS-control and the nested models fit the mild downward slope of the empirical adjustment hazard much better than the PPS-logit and Woodford specifications do.

The shaded blue bars in Figure **??** illustrate Klenow and Kryvstov's (2008) data on the average absolute price change as a function of the time since last adjustment. The size of the adjustment is largely invariant with the age of the current price, with a very mildly positive slope. Under the Woodford's hazard function, the size of the adjustment is instead strongly increasing with the time since last adjustment, since an older price is likely to be farther out of line with current costs. Under the PPS and nested specifications, the size of the adjustment varies less with the age of the price, although it is initially decreasing (due to the correction of recent large errors). It is unclear which of our specifications performs best relative to this phenomenon in the microdata.

Finally, Figure 6 illustrates the observation of Campbell and Eden (2010) that extreme prices tend to be young. The shaded blue bars represent their data, after controlling for sales; the figure shows the fraction of prices that are less than two months old, as a function of the deviation of the price from the mean price in the product group to which that price belongs. In the Campbell and Eden data, the fraction of young prices is around 50% for prices that deviate by more than 20% from the mean, whereas the fraction of young prices is only around 35% for a price equal to the mean. Extreme prices also tend to be young in the PPS and nested models; in these models extreme prices are likely to result from an extreme productivity draw compounded by an error, and are therefore unlikely to last long. However, the relation is much too strong under the PPS specification (with prices that deviate by more than 20% from the mean being around 90% young, and only 10% young prices at the mean). The nested specification shows a U-shaped relationship that is more quantitatively consistent with the data. In the Woodford specification the relationship is much flatter, though in that specification too a mild U-shape is observed.

5.3 Results: trend inflation and money supply shocks

Next, we consider the macroeconomic implications of each of our specifications concerning the effects of changes in monetary policy. Figure **??** shows how the frequency and standard deviation of price adjustment vary as trend inflation rises from 4% to 63% annually, which is the range of inflation rates documented by Gagnon (2008) for a Mexican dataset. Given this increase in the inflation rate, the frequency of price adjustment in the Mexican data increased by a factor of 1.6.[23] In the Woodford and PPS specifications, the increase in the adjustment frequency is much too high, ranging from a factor of 2.9 for Woodford-logit to 3.55 for PPS-control. The best performance comes from the nested specifications, although the change is still excessive: the frequency rises by a factor of 2.42 in Nested-logit and by 2.79 for Nested-control.

It is harder to differentiate across the models in terms of their implications for the relationship between inflation and the standard deviation of price adjustments. In some specifications, the standard deviation rises slightly with inflation; the largest rise is by a factor of 1.18 for PPS-logit. In other specifications, it falls slightly with inflation; the strongest decrease is a factor of 0.72 for PPS-control.[24] In Woodford-logit, the standard deviation is scaled down by a factor of 0.88, exactly as in the data; in Nested-control, it stays approximately constant. Thus, none of the models considered seem strongly inconsistent with this feature of the data.

Likewise, Fig. **??** shows how the fractions of positive and negative price adjustments vary with trend inflation. All models except PPS-logit perform similarly in explaining the percentage

[23]In the figure, the adjustment frequency at the low 4% inflation rate is scaled to 100 in all cases, to better compare the changes in frequency across specifications.

[24]This decrease in the standard deviation of price adjustment appears similar to that observed for the fixed menu cost specification in Costain and Nakov (2011A). In that paper, fixed menu costs implied a strongly bimodal distribution of price adjustments at a low inflation rate, which collapsed to a single-peaked distribution as inflation rose, resulting in a large decrease in the standard deviation of adjustments.

of price increass at 4% inflation (all slightly underpredict this percentage). The nested models are the best ones to explain the remaining fraction of price decreases when inflation rises: both the nested specifications imply that roughly 6% of price adjustments are negative at 63% annual inflation. For the Woodford and PPS-control specifications, on the other hand, there are only a negligible number of price decreases (0.1% or 1.1%, respectively) when inflation reaches 63%. Finally, PPS-logit, which implies noisier choice than our other specifications, still has more than 20% price decreases at a 63% inflation rate.

Figure 7 turns to the question of the effects of monetary shocks. It shows the impulse responses of inflation to consumption to a 1% money growth rate shock with monthly autocorrelation of 0.8. Somewhat surprisingly, the responses are quite similar across five of our six specifications, the exception being PPS-control. In the nested and Woodford specifications, the money supply shock leads to a fairly strong real expansion. Consumption rises by 1.8% on impact in response to a one percent money growth shock, and converges back to steady state with a half-life of roughly four months. This is a substantially less persistent response than we reported for the "smoothly state-dependent pricing" specification of Costain and Nakov (2011B), but is still a much stronger effect on consumption than would be found with a fixed menu cost model.

The Woodford specifications and nested specifications lead to almost identical impulse response functions, both for consumption and inflation, suggesting that the timing errors in Woodford's logit are the main factor responsible for the nonneutrality of the nested model too. Timing errors obviously help cause monetary nonneutrality since they imply that not all prices adjust immediately in response to a monetary shock. What is more surprising is that PPS-logit also exhibits a very similar nonneutrality. In this case, the real effects can be understood in terms of the large pricing errors implied by our estimate of the model. Given these noisy decisions, firms' adjustments may be far from optimal responses to the money supply shock. They may therefore need to readjust; note that Fig. 4 shows an adjustment hazard of almost 50% immediately after a price change for this specification. Thus firms may require several attempts before setting a satisfactory price, which slows down adjustment of the aggregate price level and leads to substantial monetary nonneutrality.

With much lower noise, the PPS-control framework behaves very differently. Errors in price setting are small, and timing is perfectly rational, so the small decision cost and risk associated with price adjustment in this specification basically act like a small menu cost. Thus, as we already noticed in Fig. 3, the PPS-control model behaves very much like a fixed menu cost model. This is true of its impulse responses too: a money supply shock causes a strong initial inflation spike, due to the immediate price changes made by the firms that cross the lower (S,s) band when the money supply increases. Thus, prices adjust quickly, and the response of consumption is correspondingly smaller, though still larger than the response reported for the fixed menu cost model in Costain and Nakov (2011B).

To see that this inflation spike is indeed a "selection effect" in the sense of Golosov and Lucas (2007), we decompose of the inflation response in Fig. 8. To construct the decomposition, define the conditionally optimal price level $p_t^{*k} \equiv \text{argmax}_p v_t(p, a^k)$, and also $x_t^{*jk} \equiv \log(p_t^{*k}/p^j)$, the desired log price adjustment of a firm at time t with productivity a^k and real price p^j. The actual log price adjustment of such a firm (call it i) can thus be decomposed as $x_{it} = x_t^{*jk} + \epsilon_{it}$, where ϵ_{it} is an error, in logs. We can then write the average desired adjustment as $\overline{x}_t^* = \sum_{j,k} x_t^{*jk} \widetilde{\Psi}_t^{jk}$, and write the fraction of firms adjusting as $\overline{\lambda}_t = \sum_{j,k} \lambda_t^{jk} \widetilde{\Psi}_t^{jk}$, and write the average log error as

$\bar{\epsilon}_t = \sum_{j,k,l} \pi_t^{lk} \log(p^l/p_t^{*k})\lambda_t^{jk}\widetilde{\Psi}_t^{jk}$. Then inflation can be written as

$$\Pi_t \ = \ \sum_{j,k} x_t^{*jk}\lambda_t^{jk}\widetilde{\Psi}_t^{jk} + \bar{\epsilon}_t. \tag{68}$$

To a first-order approximation, we can decompose the deviation in inflation at time t as

$$\Delta\Pi_t = \overline{\lambda}\Delta\overline{x}_t^* + \overline{x}^*\Delta\overline{\lambda}_t + \Delta\sum_{j,k} x_t^{jk}(\lambda_t^{jk} - \overline{\lambda}_t)\widetilde{\Psi}_t^{jk} + \Delta\bar{\epsilon}_t, \tag{69}$$

where terms without time subscripts represent steady states, and Δ represents a change relative to steady state.[25]

The "intensive margin", $\mathcal{I}_t \equiv \overline{\lambda}\Delta\overline{x}_t^*$, is the part of inflation due to changes in the average desired adjustment, holding fixed the fraction of firms adjusting. The "extensive margin", $\mathcal{E}_t \equiv \overline{x}^*\Delta\overline{\lambda}_t$, is the part due to changes in the fraction adjusting, assuming the average desired change among those who adjust equals the steady-state average in the whole population. The "selection effect", $\mathcal{S}_t \equiv \Delta\sum_{j,k} x_t^{jk}(\lambda_t^{jk} - \overline{\lambda}_t)\widetilde{\Psi}_t^{jk}$, is the inflation caused by redistributing adjustment opportunities from firms desiring small price adjustments to firms desiring large adjustments, while fixing the total number adjusting. The last term, $\Delta\bar{\epsilon}_t$, is the change in the average log error. Figure 8 reports the first three components of the inflation decomposition for our six specifications. We see that indeed, the spike of inflation on impact in PPS-control is a selection effect. Interestingly, the majority of the inflation response is also attributed to the selection component in the nested specifications, but this selection effect occurs gradually over time. The intensive margin is much smaller, and the extensive margin is negligible, in all the specifications considered.

Finally, in Table 4, we assess the degree of nonneutrality in our model by running two calculations from Golosov and Lucas (2007). Assuming for concreteness that money shocks are the only cause of macroeconomic fluctuations, we calibrate the standard deviation of the money shock for each specification to perfectly match the standard deviation of quarterly inflation (one quarter of one percent) in US data. We then check what fraction of the time variation in US output growth can be explained by those shocks. In the Woodford and nested specifications, these money shocks would explain around 80% or 90% of the observed variation in US output growth. PPS-logit would explain 67% of output growth variation, while PPS-control would explain much less, only 38%, consistent with the strong inflation spike and small output response observed in Fig. 7 for this specification. In the last line of the table, we also report "Phillips curve" coefficients, that is, estimates from an instrumental variables estimate of the effect of inflation on output, instrumenting inflation by the exogenous money supply process. The coefficient is more than twice as large for the nested, Woodford, and PPS-logit cases than it is for PPS-control.

6 Conclusions

This paper has modeled nominal price rigidity as a near-rational phenomenon. Price adjustment is costly, but the interpretation of the costs is not the usual one: they represent the cost of decision-making by management.

[25] See Costain and Nakov (2011B) for further discussion of this decomposition.

We operationalize this idea by adopting a common assumption from game theory: a "control cost" function that depends on the precision of the decision. Following Mattsson and Weibull (2002), we assume that precision is measured by relative entropy, and then show that decisions are random variables with logit form. This standard game-theoretic result is directly applicable to the question of *which* price to set when the firm has decided to make an adjustment. We show how to extend this result to model the decision of *when* to adjust the price. Just as the cost of the price choice is assumed proportional to relative entropy compared with a uniform price distribution, the cost of the timing choice is assumed proportional to the relative entropy of the adjustment hazard, compared with a uniform adjustment hazard. The resulting model of near rational choice has just two parameters: a noise parameter measuring the accuracy of decisions, and a rate parameter measuring the speed of decisions.

We shut down the errors on each choice margin– the timing margin and the pricing margin– to see the role played by each type of error. The model with pricing errors, but perfect adjustment timing, implies that prices are sticky when they are near the optimum, because of the risk of choosing a worse price; therefore we call this specification "precautionary price stickiness" (PPS). This special case has only one free parameter– the degree of noise in the pricing decision. Our simulations show that noise in the pricing decision helps match a variety of features of the price adjustment microdata, but with only one free parameter the model cannot in general match both the typical size of adjustment and its typical frequency. We refer to the model with errors in timing but perfect pricing decisions as "Woodford's logit", because the functional form for the adjustment hazard is the same weighted logit derived by Woodford (2009) for a rational inattention model. Both the Woodford specification, and our general nested specification, have two free parameters: the decision accuracy and the decision rate. But with a few exceptions the nested specification fits the data far better than the Woodford specification does.

With just two parameters, the nested specification fits well both the timing and size of price adjustments. As microdata show, both large and small price adjustments coexist in the distribution. The adjustment hazard is largely flat, with a mild downward slope. Extreme prices are more likely to have been recently set. The nested model is well-behaved as inflation rises from 4% to 63% annually, and it performs better than the PPS or Woodford specifications in describing how the distribution of price adjustments changes with inflation (in the light of Gagnon's (2008) Mexican data). Both the nested model and the Woodford model imply a realistic degree of monetary nonneutrality in response to money growth shocks (though substantially less than a Calvo model with the same average rate). While this paper has focused on comparing the nested specification with the PPS and Woodford specifications, when comparing our calculations of the nested specification in this paper with our previous calculations of the Calvo model and fixed menu cost model in other papers we see that it outperforms those models too.

In recent literature a variety of models have been proposed to explain puzzles from microeconomic pricing data; meanwhile, empirical work continues to discover more puzzles. While we do not claim to explain all features of the data, we find it encouraging that our sparsely parameterized model works as well as it does, and that it can be incorporated into a macroeconomic model in a tractable way. We intend to extend this model of intermittent adjustment to other types of decision problems in future work.

References

[1] Akerlof, George, and Janet Yellen (1985), "A near-rational model of the business cycle with wage and price inertia." *Quarterly Journal of Economics* 100 (Supplement), pp. 823-38.

[2] Álvarez, Luis (2008), "What do micro price data tell us on the validity of the New Keynesian Phillips curve?" *Economics: The Open-Access, Open-Assessment E-Journal* 2 (19).

[3] Álvarez, Fernando; Francesco Lippi, and Luigi Paciello (2011), "Optimal price setting with observation and menu costs." *Quarterly Journal of Economics.*

[4] Álvarez, Fernando; Martín González-Rozada; Andy Neumeyer; and Martín Beraja (2011), "From hyperinflation to stable prices: Argentina's evidence on menu cost models." Manuscript, Univ. of Chicago.

[5] Anderson, Simon; Jacob Goeree, and Charles Holt (2002), "The logit equilibrium: a perspective on intuitive behavioral anomalies." *Southern Economic Journal* 69 (1), pp. 21-47.

[6] Ash, Robert (1965), *Information Theory.* Dover Publications.

[7] Baron, Richard, Jaques Durieu, Hans Haller, and Philippe Solal (2002), "Control costs and potential functions for spatial games." *International Journal of Game Theory* 31, pp. 541-561.

[8] Barro, Robert (1972), "A theory of monopolistic price adjustment." *Review of Economic Studies* 39 (1), pp. 17-26.

[9] Caballero, Ricardo, and Eduardo Engel (1993), "Microeconomic rigidities and aggregate price dynamics." *European Economic Review* 37 (4), pp. 697-711.

[10] Caballero, Ricardo, and Eduardo Engel (1999), "Explaining investment dynamics in U.S. manufacturing: a generalized (S,s) approach." *Econometrica* 67 (4), pp. 783-862.

[11] Caballero, Ricardo, and Eduardo Engel (2007), "Price stickiness in Ss models: new interpretations of old results." *Journal of Monetary Economics* 54, pp. 100-21.

[12] Calvo, Guillermo (1983), "Staggered prices in a utility-maximizing framework." *Journal of Monetary Economics* 12, pp. 383-98.

[13] Campbell, Jeffrey, and Benjamin Eden (2010), "Rigid prices: evidence from U.S. scanner data."FRB Chicago WP2005-08 (revised July 2010).

[14] Caplin, Andrew, and Daniel Spulber (1987), "Menu costs and the neutrality of money." *Quarterly Journal of Economics* 102 (4), pp. 703-25.

[15] Caplin, Andrew, and John Leahy (1991), "State-dependent pricing and the dynamics of money and output." *Quarterly Journal of Economics* 106 (3), pp. 683-708.

[16] Christiano, Lawrence; Martin Eichenbaum, and Charles Evans (1999), "Monetary policy shocks: what have we learned and to what end?" In J. Taylor and M. Woodford, eds. (1999), *Handbook of Macroeconomics*, Volume 1A. Elsevier.

[17] Costain, James, and Anton Nakov (2011A), "Price adjustments in a general model of state-dependent pricing." *Journal of Money, Credit and Banking* 43 (2-3), pp. 385-406.

[18] Costain, James, and Anton Nakov (2011B), "Distributional dynamics under smoothly state-dependent pricing." *Journal of Monetary Economics* 58 (6-8), pp. 646-665.

[19] Costain, James, and Anton Nakov (2011C), "Precautionary price stickiness." European Central Bank Working Paper #1375.

[20] Cover, Thomas, and Joy Thomas (2006), *Elements of Information Theory*, 2nd ed. Wiley Interscience.

[21] Danziger, Leif (1999), "A dynamic economy with costly price adjustments." *American Economic Review* 89 (4), pp. 878-901.

[22] Dotsey, Michael; Robert King; and Alexander Wolman (2008), "Inflation and real activity with firm-level productivity shocks." Manuscript, Boston University.

[23] Eichenbaum, M.; N. Jaimovich, and S. Rebelo (2011), "Reference prices, costs and nominal rigidities." *American Economic Review* 101 (1), pp. 234-262.

[24] Gagnon, Etienne (2007), "Price setting under low and high inflation: evidence from Mexico." *Quarterly Journal of Economics* 124, pp. 1221-1263.

[25] Gautier, Erwan, and Ronan Le Saout (2012), "The dynamics of gasoline prices: evidence from daily French micro data." Document de Travail #375, Banque de France.

[26] Gertler, Mark, and John Leahy (2008), "A Phillips curve with an Ss foundation." *Journal of Political Economy* 116 (3), pp.533-72.

[27] Golosov, Mikhail, and Robert E. Lucas, Jr. (2007), "Menu costs and Phillips curves." *Journal of Political Economy* 115 (2), pp. 171-99.

[28] Gorodnichenko, Yuriy (2008), "Endogenous information, menu costs, and inflation persistence." NBER Working Paper #14184.

[29] Guimaraes, and Kevin Sheedy (2011), "Sales and monetary policy." *American Economic Review* 101 (2), pp. 844-876.

[30] Hellwig, Christian, and Venky Venkataswaran (2009), "Setting the right prices for the wrong reasons." *Journal of Monetary Economics* 56, Supplement, pp. 57-77.

[31] Kehoe, Patrick, and Virgiliu Midrigan (2010), "Prices are sticky after all." NBER Working Paper #16364.

[32] Kiley, Michael (2002), "Partial adjustment and staggered price setting." *Journal of Money, Credit and Banking* 34 (2), pp. 283-98.

[33] Klenow, Peter, and Oleksiy Kryvtsov (2008), "State-dependent or time-dependent pricing: does it matter for recent US inflation?", *Quarterly Journal of Economics* 123, pp. 863-904.

[34] Klenow, Peter, and Benjamin Malin (2010), "Microeconomic evidence on price setting." In *Handbook of Monetary Economics*, Benjamin Friedman and Michael Woodford, eds., vol. 3, number 3.

[35] Mankiw, N. Gregory (1985), "Small menu costs and large business cycles: a macroeconomic model of monopoly." *Quarterly Journal of Economics* 100 (2), pp. 529-37.

[36] Marsili, Matteo (1999), "On the multinomial logit model." *Physica A* 269, pp. 9-15.

[37] Matejka, Filip, and Alisdair McKay (2011), "Rational inattention to discrete choices: a new foundation for the multinomial logit model." Manuscript, CERGE-EI.

[38] Matejka, Filip (2011), "Rationally inattentive seller: sales and discrete pricing." Manuscript, CERGE-EI.

[39] Mattsson, Lars-Goram, and Jorgen Weibull (2002), "Probabilistic choice and procedurally bounded rationality." *Games and Economic Behavior* 41, pp. 61-78.

[40] McKelvey, Richard, and Thomas Palfrey (1995), "Quantal response equilibrium for normal form games." *Games and Economic Behavior* 10, pp. 6-38.

[41] McKelvey, Richard, and Thomas Palfrey (1998), "Quantal response equilibrium for extensive form games." *Experimental Economics* 1, pp. 9-41.

[42] Midrigan, Virgiliu (2011), "Menu Costs, Multi-Product Firms and Aggregate Fluctuations." *Econometrica* 79 (4), pp. 1139-1180.

[43] Nakamura, Emi, and Jón Steinsson (2008), "Five facts about prices: a reevaluation of menu cost models." *Quarterly Journal of Economics* 123 (4), pp. 1415-1464.

[44] Nakamura, Emi, and Jón Steinsson (2011), "Price setting in forward-looking customer markets." *Journal of Monetary Economics* 58, pp. 220-233.

[45] Reiter, Michael (2009), "Solving heterogeneous-agent models by projection and perturbation." *Journal of Economic Dynamics and Control* 33 (3), pp. 649-65.

[46] Shannon, Claude (1948), "A mathematical theory of communication."

[47] Sheshinski, Eytan, and Yoram Weiss (1977), "Inflation and costs of price adjustment." *Review of Economic Studies* 44 (2), pp. 287-303.

[48] Sims, Christopher (1998), "Stickiness." *Carnegie-Rochester Conference Series on Public Policy* 49, pp. 317-56.

[49] Sims, Christopher (2003), "Implications of rational inattention." *Journal of Monetary Economics* 50, pp. 665-90.

[50] Stahl, Dale (1990), "Entropy control costs and entropic equilibrium. " *International Journal of Game Theory* 19, pp. 129-138.

[51] Van Damme, Eric (1991), *Stability and Perfection of Nash Equilibrium*, 2nd edition. Springer Verlag.

[52] Woodford, Michael (2008), "Information-constrained state-dependent pricing." *Journal of Monetary Economics* 56, pp. S100-S124.

[53] Zbaracki, Mark; Mark Ritson; Daniel Levy; Shantanu Dutta; and Mark Bergen (2004), "Managerial and customer costs of price adjustment: direct evidence from industrial markets." *Review of Economics and Statistics* 86 (2), pp. 514-33.

Table 1: Adjustment parameters.

	Woodford logit	Woodford control	PPS logit	PPS control	Nested logit	Nested control
$\bar{\lambda}$	0.044	0.045	–	–	0.083	0.22
κ_π	–	–	0.049	0.0044	0.013	0.018
κ_λ	0.0051	0.0080	–	–	0.013	0.018

Table 2: Model-Simulated Statistics and Evidence (1% annual inflation)

	Woodford logit	Woodford control	PPS logit	PPS control	Nested logit	Nested control	Data
Freq. of price changes	10.2	10.2	10.2	10.2	10.2	10.2	10.2
Mean absolute price change	4.88	4.68	14.0	6.72	8.11	7.51	9.90
Std of price changes	5.51	5.27	17.0	7.32	10.1	9.30	13.2
Kurtosis of price changes	2.24	2.22	2.58	2.37	3.48	3.40	4.81
Percent of price increases	62.7	63.3	55.2	62.3	58.3	58.8	65.1
Percent of changes $\leq 5\%$	47.9	49.7	16.5	27.9	31.5	33.6	35.4
Pricing costs*	0	0	0	0.174	0	0.509	
Timing costs*	0	0.167	0	0	0	0.361	
Loss relative to full rationality*	0.258	0.416	0.665	0.365	0.582	1.41	

Note: All statistics refer to regular consumer price changes excluding sales, and are stated in percent.
Quantities with an asterisk are stated as a percentage of monthly average revenues.
Dataset: Dominick's.

Table 3: Model-Simulated Statistics and Evidence (63% annual inflation)

	Woodford logit	Woodford control	PPS logit	PPS control	Nested logit	Nested control	Data
Freq. of price changes	29.0	29.6	32.1	35.5	24.2	26.0	
Mean absolute price change	14.0	13.7	19.7	11.5	17.9	16.6	
Std of price changes	4.87	4.86	20.1	5.26	11.7	10.9	
Kurtosis of price changes	3.14	3.12	3.32	6.38	4.64	4.43	
Percent of price increases	99.9	99.9	78.5	98.9	93.3	93.1	
Percent of changes $\leq 5\%$	4.4	4.4	11.2	7.71	7.83	8.71	
Pricing costs*	0	0	0	0.557	0	1.06	
Timing costs*	0	1.00	0	0	0	0.66	
Loss relative to full rationality*	0.752	1.65	1.92	1.00	1.72	3.25	

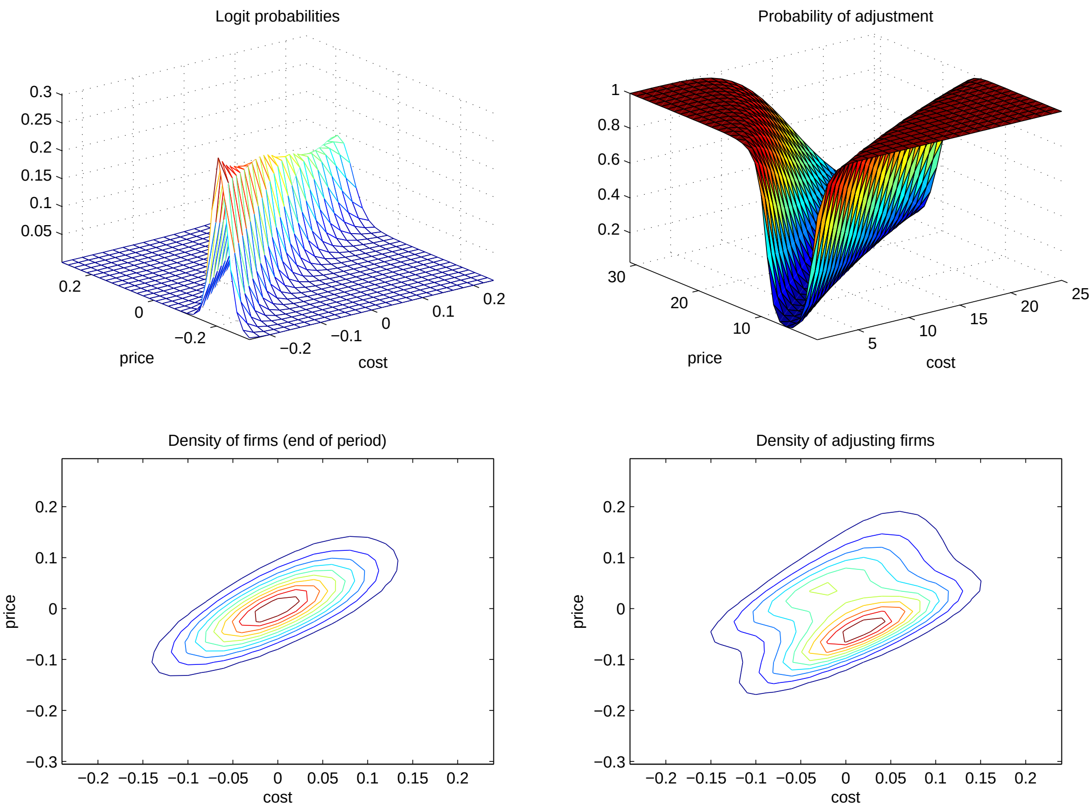
Note: All statistics refer to regular consumer price changes excluding sales, and are stated in percent.
Quantities with an asterisk are stated as a percentage of monthly average revenues.
Dataset: Gagnon (2008) Mexican data.

Table 4: Variance decomposition and Phillips curves

Correlated money growth shock ($\phi_z = 0.8$)	Woodford logit	Woodford control	PPS logit	PPS control	Nested logit	Nested control	Data*
Freq. of price changes (%)	10.2	10.2	10.2	10.2	10.2	10.2	10.2
Std of money shock (%)	0.16	0.16	0.16	0.12	0.17	0.16	
Std of qtrly inflation (%)	0.25	0.25	0.25	0.25	0.25	0.25	0.25
% explained by μ shock alone	100	100	100	100	100	100	
Std of qtrly output growth (%)	0.41	0.41	0.34	0.20	0.45	0.40	0.51
% explained by μ shock alone	80	81	67	38	89	79	
Slope coeff. of Phillips curve*	0.32	0.33	0.31	0.15	0.38	0.33	
R^2 of regression	0.96	0.97	0.999	0.85	0.99	0.97	

*The "slope coefficients" are 2SLS estimates of the effect of inflation on consumption
First stage: $\pi_t^q = \alpha_1 + \alpha_2 \mu_t^q + \epsilon_t$; second stage: $c_t^q = \beta_1 + \beta_2 \hat{\pi}_t^q + \varepsilon_t$, where the instrument μ_t^q is the exogenous growth rate of the money supply and the superscript q indicates quarterly averages.
Dataset: Dominick's.

Figure 1: Price change distributions and adjustment function: nested control cost model.



Notes:

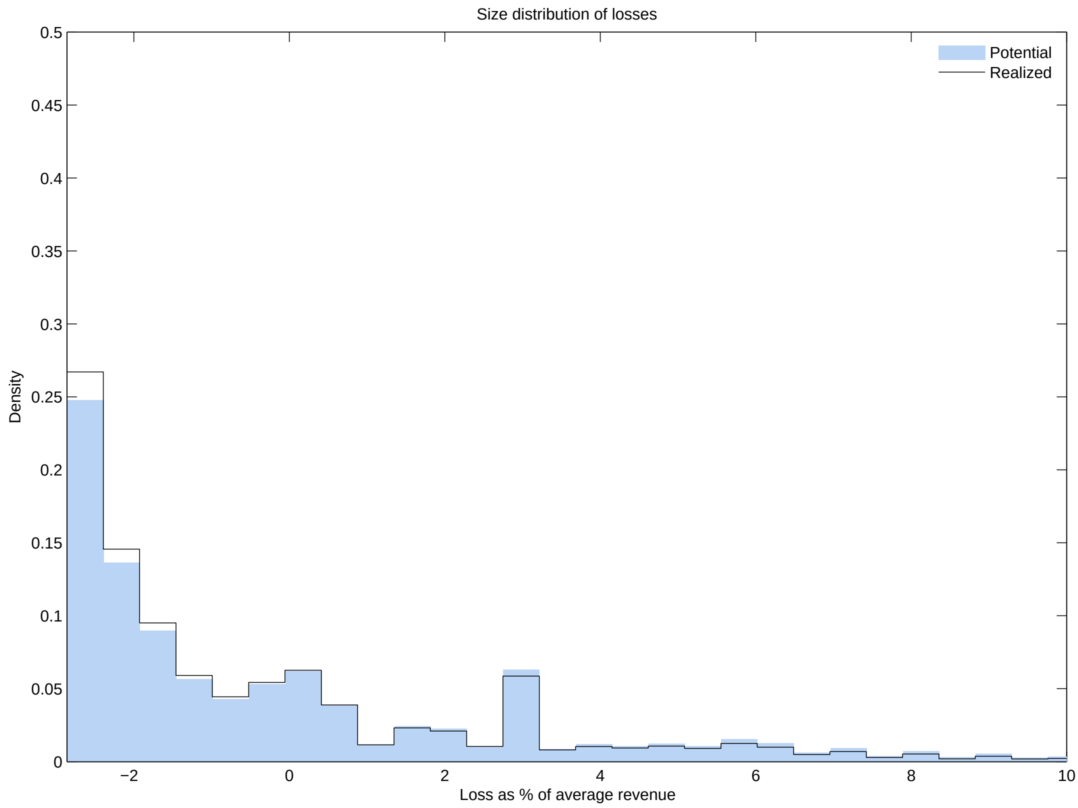
First panel: price distribution conditional on cost.

Second panel: adjustment probability conditional on price and cost.

Third panel: contour plot of density of firms at time of production.

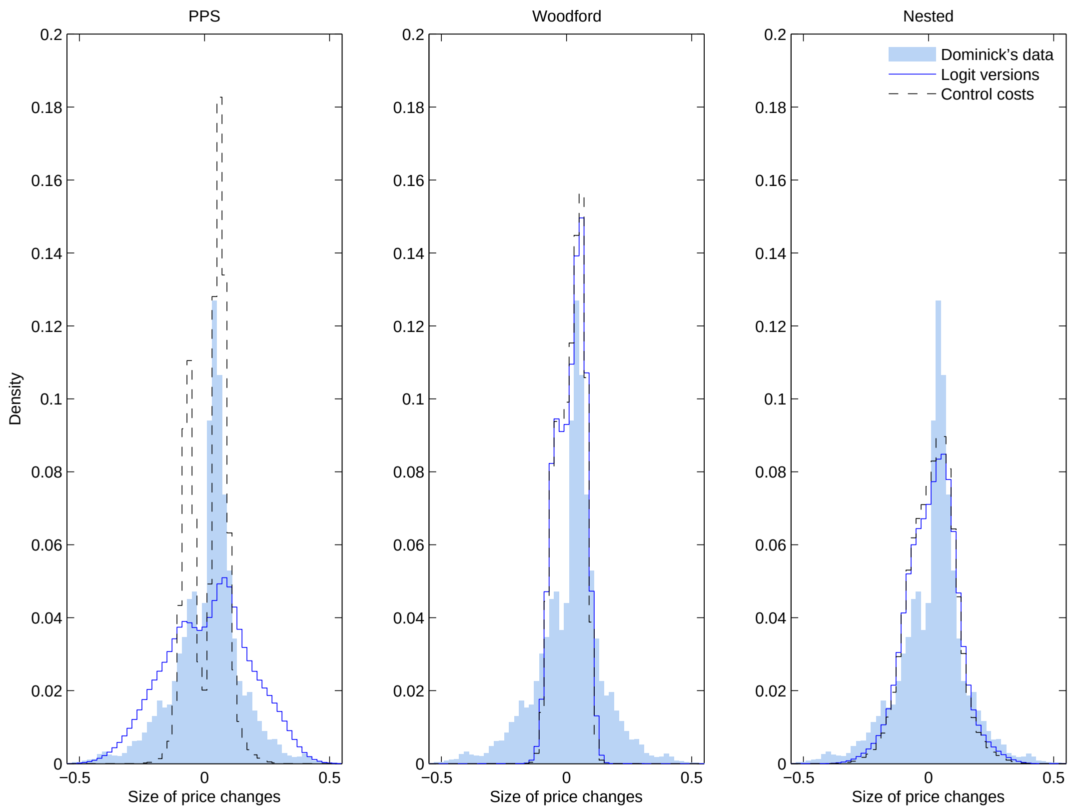
Fourth panel: contour plot of density of adjusting firms.

Figure 2: Losses from failure to adjust: nested control cost model.



Notes:
Loss from not adjusting, expressed as a percentage of average revenues. Potential losses before adjustments occur (shaded blue) and realized losses after adjustments (black line).

Figure 3: Distribution of price adjustents: comparing models.



Notes:

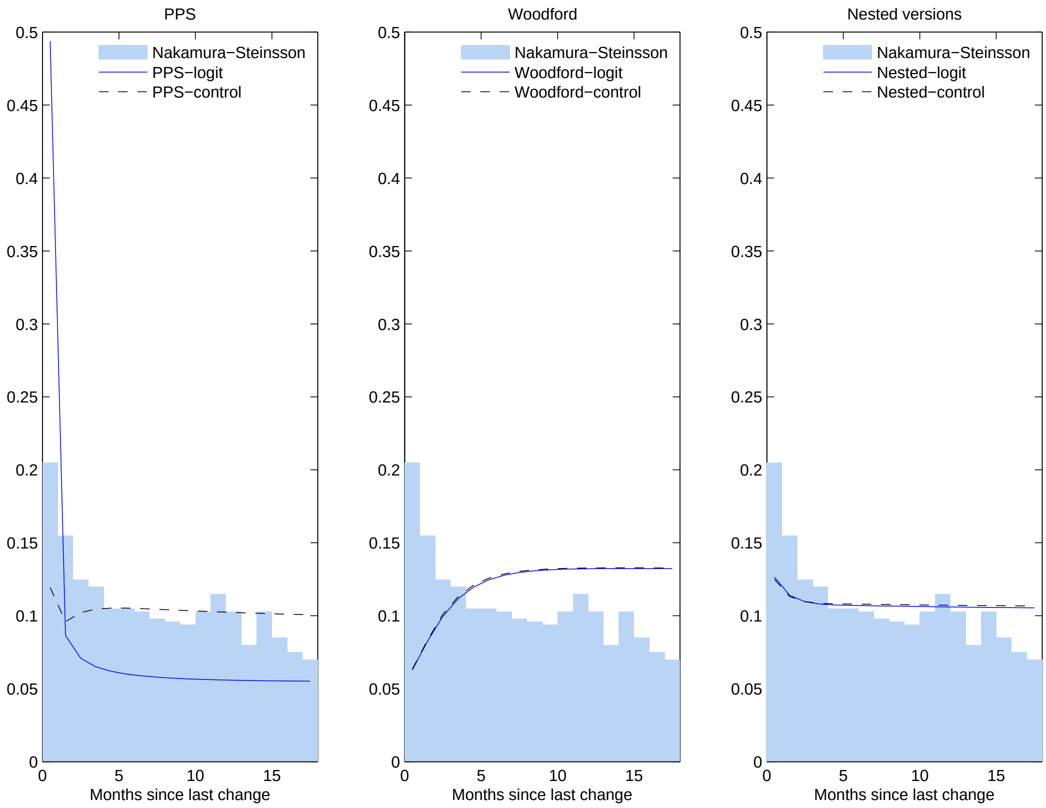
Comparing histogram of price adjustments across models.

Shaded area: histogram of price adjustments in Dominick's data.

Solid lines: histograms of price adjustments in logit versions of the model.

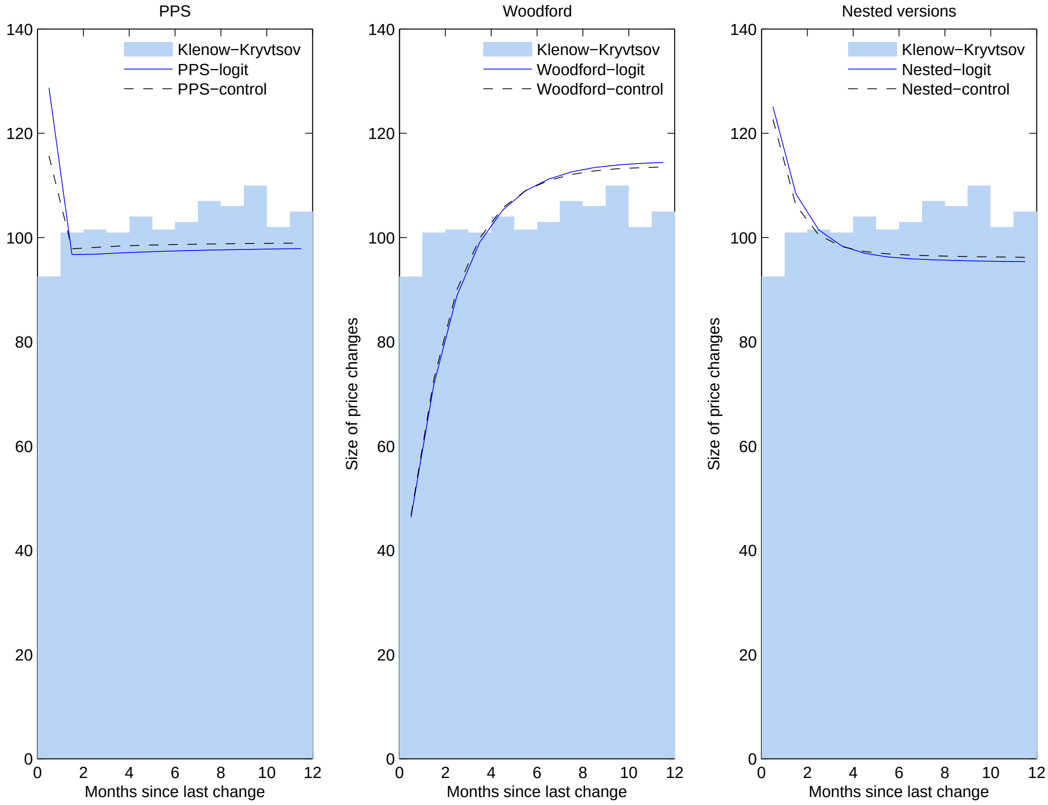
Dashed lines: histograms of price adjustments in control cost versions of the model.

Figure 4: Price adjustment hazard: comparing models.



Notes:
Adjustment probability as a function of time since last price change.
Shaded area: price adjustment hazard in Dominick's data.
Solid lines: price adjustment hazards in logit versions of the model.
Dashed lines: price adjustment hazards in control cost versions of the model.

Figure 5: Mean adjustment and price duration: comparing models.



Notes:

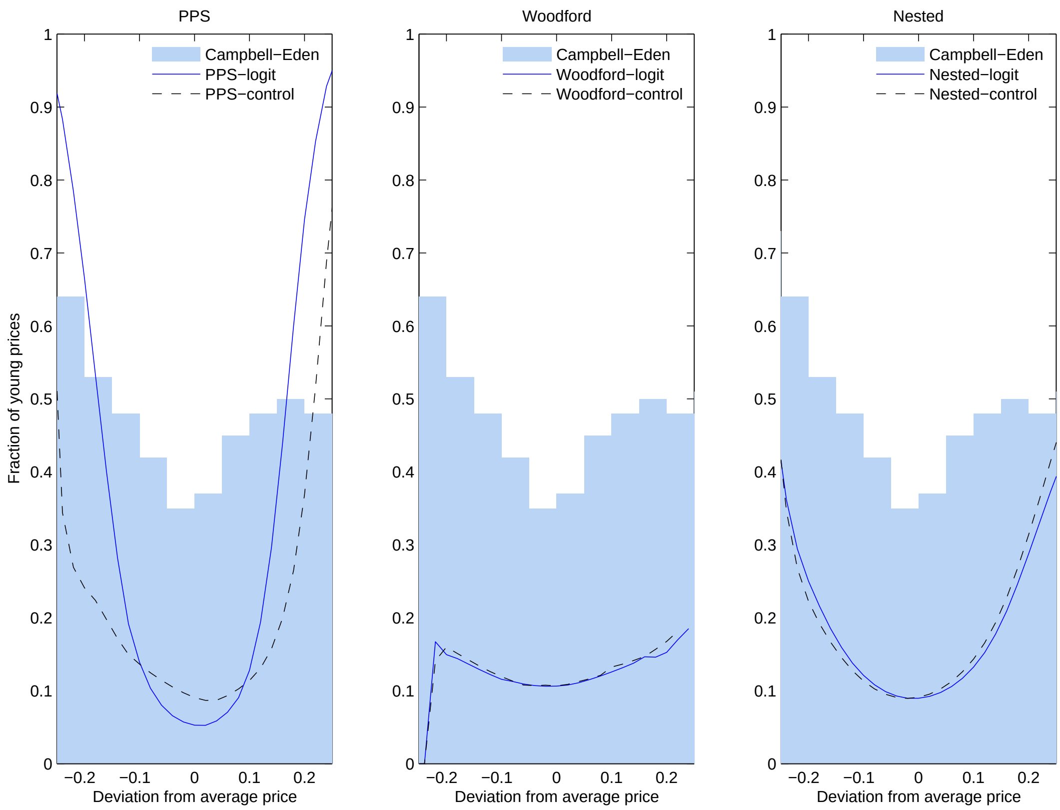
Mean absolute size of price adjustment as function of time since last price change.

Shaded area: Klenow-Kryvstov dataset.

Solid lines: logit versions of the model.

Dashed lines: control cost versions of the model.

Figure 6: Extreme prices tend to be young: comparing models.



Notes:

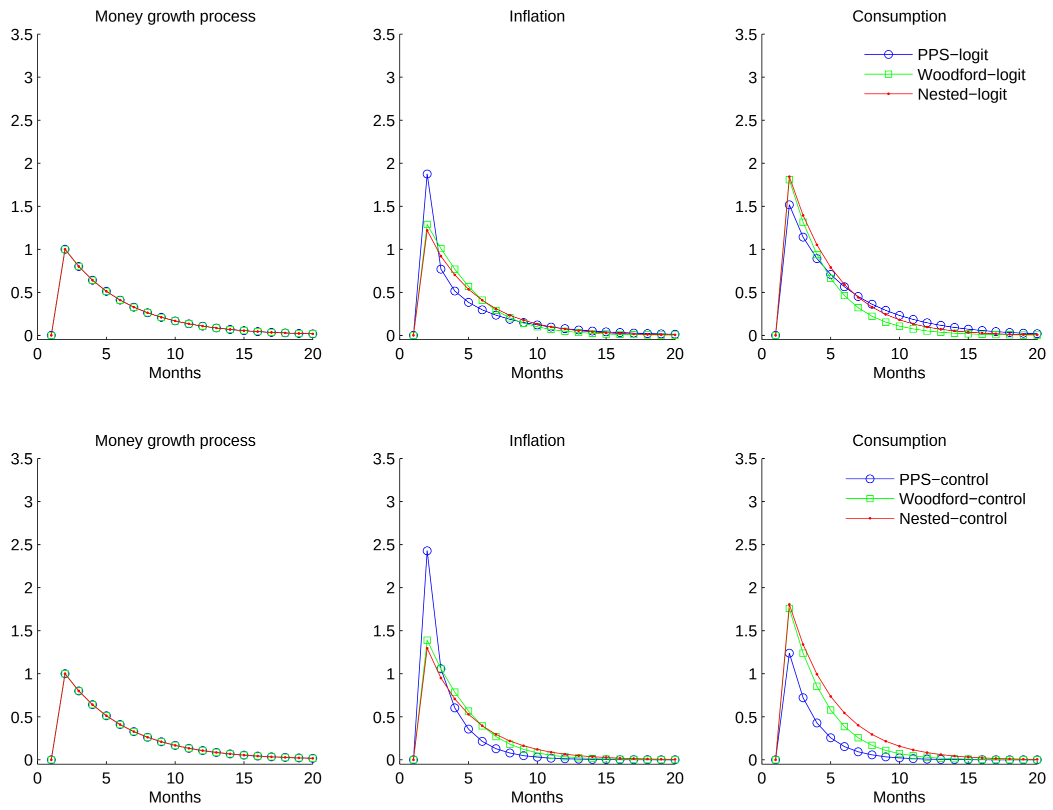
Fraction of prices set within the last two months, as a function of deviation from average price in product class.

Shaded area: Campbell-Eden dataset.

Solid lines: logit versions of the model.

Dashed lines: control cost versions of the model.

Figure 7: Impulse responses to money growth shock: comparing models.



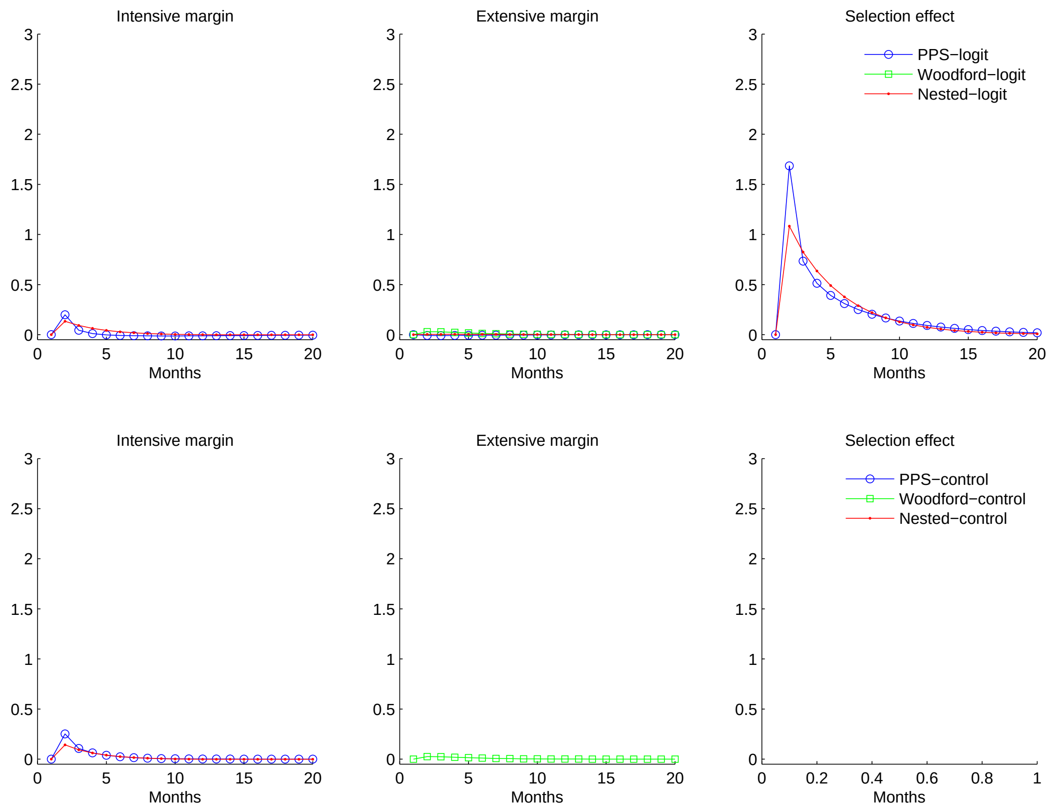
Notes:

Top row: impulse responses of inflation and consumption to money growth shock with autocorrelation 0.8 (monthly), logit specifications.

Bottom row: impulse responses of inflation and consumption to money growth shock with autocorrelation 0.8 (monthly), control cost specifications.

Green lines with circles: PPS versions. Blue lines with diamonds: Wood versions. Red lines: Nested versions.

Figure 8: Decomposition of inflation impulse responses: comparing models.



Notes:

Top row: decomposition of inflation impulse response to money growth shock with autocorrelation 0.8 (monthly), logit specifications.

Bottom row: decomposition of inflation impulse response to money growth shock with autocorrelation 0.8 (monthly), control cost specifications.

Green lines with circles: PPS versions. Blue lines with diamonds: Wood versions. Red lines: Nested versions.