Logit price dynamics

JAMES COSTAIN*

ANTON NAKOV†

August 2018

Abstract

We model retail price stickiness as the result of costly, error-prone decision-making. Under our assumed cost function for the precision of choice, the timing of price adjustments and the prices firms set are both logit random variables. Errors in the prices firms set help explain micro facts related to the size of price changes, the behavior of adjustment hazards, and the variability of prices and costs. Errors in adjustment timing increase the real effects of monetary shocks, by reducing the "selection effect". Allowing for both types of errors also helps explain how trend inflation affects price adjustment.

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

JEL Codes: E31, D81, C73

*JAMES COSTAIN is an Economist at the Banco de España and the European Central Bank (E-mail: james.costain@ecb.europa.eu).

†ANTON NAKOV is an Economist at the European Central Bank and an Affiliate of the Centre for Economic Policy Research (E-mail: anton.nakov@ecb.europa.eu).

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). In theory, even small menu costs might impede price adjustment sufficiently to alter aggregate dynamics in a significant way (Mankiw 1985). But quantitatively, Golosov and Lucas (2007) showed that fixed menu costs do little to generate aggregate price stickiness in a macroeconomic model with realistically large firm-specific shocks. The dynamics of their model are quite close to monetary neutrality, so fixed menu costs seem unpromising to explain the nontrivial real effects of monetary shocks observed in macroeconomic data (*e.g.* Christiano, Eichenbaum, and Evans, 1999). Moreover, detailed microeconomic evidence suggests that menu 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 related to information processing and decision making. This raises the question: can costs of decision-making explain micro and macro evidence on price dynamics better than fixed menu costs do? And how exactly might these costs be modeled?

This paper proposes a simple model of price stickiness based on costly choice, estimates two of its free parameters, and shows by simulation that this suffices for consistency with a wide variety of microeconomic and macroeconomic evidence. Two key considerations motivate our setup. First, if choice is costly, then decisions are likely to be error-prone. This makes it natural to treat decision outcomes as random variables, instead of treating actions as deterministic. Second, it is natural to assume that greater precision is costly — for example, by taking more time to consider more relevant variables, a decision maker might increase the probability of selecting the best feasible option. Motivated by these points, we adopt the "control cost" approach from game theory (see, for example, van Damme 1991). Formally, instead of modeling the choice of an optimal action directly, this approach defines the decision problem as the choice of a probability distribution over possible actions.[2] The problem is constrained by a cost function under which more precise decisions (more concentrated distributions) are more expensive. Making any given decision in a perfectly precise way is feasible, but is usually not worth the cost. Therefore actions are random variables correlated with fundamentals, instead of being deterministic functions of fundamentals.

In the context of dynamic price setting, a firm faces choices on two key margins: *when* to change the price of a product it sells, and *what new price* to set. In contrast to existing game theoretic applications of control costs, we allow for errors on both these margins. The resulting distribution of errors depends on the functional form of the control costs. It is especially convenient to measure precision in terms of entropy, defining costs as a linear function of the relative entropy between the distribution of actions and an exogenous default distribution. Under this cost function, the distribution of actions is a multinomial logit. This implies that the probability of taking any given action increases smoothly with the value of that action, compared with other feasible actions. General equilibrium then takes the form of a logit equilibrium:[3] each decision maker plays a logit in which the values of actions are evaluated assuming that others' choices are logits too. The decision costs backed out from our benchmark calibration do not seem excessive: firms spend roughly 0.9% of revenue on decision-making, and in addition incur a loss of roughly 0.5% of revenue due to suboptimal choices.

Many papers have shown that an entropy-based cost function can "microfound" a logit distribution of actions (Stahl 1990; Marsili 1999; Mattsson and Weibull 2002; Bono and Wolpert 2009; Matejka and McKay 2015).[4] But previous studies have typically focused on decisions taken at known, exogenously given points in time. We extend the logit model to apply to contexts of intermittent adjustment where a key decision is *when* changes should occur, showing that if the cost of choosing a time-varying adjustment hazard depends linearly on its relative entropy, relative to a constant hazard, then the chosen adjustment hazard takes the form of a binary logit. In addition to the noise parameter that governs the *accuracy* of choice in a standard logit model, our dynamic setup also features a parameter related to the *speed* of choice. The inclusion of the speed parameter ensures that our model has a well-defined continuous-time limit, and thus clarifies how parameters must be adjusted if the frequency of the data or of the simulation is changed.

While both the size and the timing of price adjustments are plausibly subject to error, we run simulations that shut down one type of mistakes or the other in order to see what each one contributes empirically. We find that errors in the *size* of price changes help explain some "puzzling" aspects of retail price microdata. In particular, unlike a fixed menu cost model, our setup implies that many large and small price changes coexist (Klenow and Kryvstov 2008; Midrigan 2011; Klenow and Malin 2010, "Fact 7"). It implies that the adjustment hazard is nearly flat, but slightly decreasing in the first few months, as found by empirical work that controls for heterogeneity in hazards (Nakamura and Steinsson 2008, "Fact 5"; Klenow and Malin 2010, "Fact 10"). Likewise, we find that the size of price changes is largely independent of the time since last adjustment (Klenow and Malin 2010, "Fact 10"). Many alternative models, including the Calvo (1983) model, instead imply that adjustments get larger, quickly, as the time since the previous change increases. Also, we find that the highest and lowest prices are more likely to have been set recently than prices near the center of the distribution (Campbell and Eden 2014). Finally, prices are more volatile than costs, as documented by Eichenbaum, Jaimovich, and Rebelo (2011), whereas the opposite is true in both the Calvo and fixed menu cost models.

While errors in the size of price adjustments help reproduce patterns in microdata, they do not by themselves yield strong real effects of monetary policy. Indeed, in our estimate of the specification with pricing errors only, the real effects are just as small as those in a fixed menu cost model. But as long as we allow for mistakes in the *timing* of price adjustments, monetary nonneutrality increases substantially. The cause of the nonneutrality is the same as in the Calvo framework: by weakening the correlation between the value of adjustment and the probability of adjustment, the “selection effect” highlighted by Caplin and Spulber (1987) and Golosov and Lucas (2007) is reduced. Thus the adjustment process combines state-dependent and time-dependent features (consistent with empirical evidence of Klenow and Kryvtsov, 2008), and the degree of nonneutrality lies roughly halfway between that of the fixed menu cost specification and that of the Calvo model. In contrast with the Calvo setup, our model also does a good job in reproducing the effects of trend inflation on price adjustment— particularly the effects on the typical size of price changes, and on the fraction of changes that are increases, which are margins where the fixed menu cost model performs poorly. The presence of both types of errors is crucial for our model’s fit to these effects of trend inflation.

1.1 Related literature

A wave of recent research has documented intermittent price adjustment in new databases from the retail sector, including work by Klenow and Kryvtsov (2008), Nakamura and Steinsson (2008), Klenow and Malin (2010), and Eichenbaum, Jaimovich, and Rebelo (2011). In response, macroeconomists have built numerical models of pricing under fixed or stochastic menu costs with both aggregate and firm-specific shocks, fitting them to microdata and then studying their macroeconomic implications; key papers include Golosov and Lucas (2007); Kehoe and Midrigan (2015); Álvarez, Beraja, González, and Neumeyer (2013); and Dotsey, King, and Wolman (2013). There have also been numerous proposals of other mechanisms, beyond the simple menu cost framework, that may better match microdata. Midrigan (2011) proposes that the firm may pay a single fixed cost to change the prices of several products. Matejka (2016) and Stevens (2015) show that when information flow is costly to the firm, retail prices may fluctuate between a few discrete values, generating some price changes that look like temporary “sales”, a phenomenon we will not address here. Álvarez, Lippi, and Paciello (2011) and Demery (2012) assume there is a fixed cost to obtain information as well as a fixed cost to adjust the price; like our own framework, these “menu cost and observation cost” models match many empirical facts while requiring just two free parameters to model the adjustment process.

Our paper contrasts with other recent work by exploring a different mechanism for nominal rigidity: errors derived from costly decision-making.[5,6] While allowing for errors may be unusual in macroeconomics, it is central to microeconometrics (though econometric "error terms" are not always interpreted as mistakes). In representative-agent macro models, ignoring errors is not necessarily inconsistent with microeconometrics, since to a first approximation, errors might cancel out. But when fitting a heterogeneous-agent macroeconomic model to the full distribution of adjustments in microdata, that argument does not apply: if there are any errors at all, they are likely to increase the variance of observed adjustments, so a calibration without errors would (for example) mistakenly overestimate the variance of the underlying exogenous shocks. In this sense, the microdata-based calibration strategies in many recent state-dependent pricing papers may be further removed from standard practice in micro- and macroeconomics than our model is.

The logit equilibrium framework has been influential in experimental game theory, because allowing for errors helps explain play in many contexts where Nash equilibrium performs poorly (Anderson, Goeree, and Holt 2002), but we are unaware of any dynamic general equilibrium macroeconomic models based on logit equilibrium, prior to our own work.[7] While McKelvey and Palfrey defined logit equilibrium both for normal form (1995) and extensive form (1998) games, we had to extend their framework in order to model errors in the *timing* of price changes. Our setup applies the same logic to timing decisions that it applies on the pricing margin. In a static context, logit choice is derived by penalizing the entropy of the random choice, relative to a fixed default distribution. Likewise, we derive a weighted binary logit for adjustment timing by penalizing the entropy of the random time of adjustment, relative to a constant default hazard. Our hazard has the same functional form derived by Woodford (2008), though his microfoundations differ: he assumes firms face a constraint on information flow, plus a fixed cost of purchasing full information.

Woodford's (2008, 2009) papers form part of the "rational inattention" literature initiated by Sims (2003), where economic agents face costs associated with information flow. Our approach is closely related, but distinct: we assume the firm has all the information it needs to make an optimal decision, but that doing the required calculation precisely is costly.[8] For example, greater precision might require managers to consider more payoff-relevant variables, or higher-order terms, which might take more time. Some recent rational inattention papers provide further motivation for our approach. Khaw, Stevens, and Woodford (2017) present laboratory evidence for decision costs above and beyond the costs of obtaining information: players in their experiment reset the control variable less frequently and more noisily than Bayesian-rational decision-makers would, given the information they have. Steiner, Stewart, and Matejka (2017) show that a dynamic rational inattention problem is equivalent to a control cost model with an optimally-chosen default distribution. While our paper instead sets the default distribution in the control cost function exogenously, this has an important practical advantage: it dramatically reduces the dimensionality of our calculations. This is because choices in a rational inattention model are conditioned on a prior (typically a high-dimensional object), whereas in our setup, choices are just conditioned on the true state of the world. This makes dynamic general equilibrium modeling tractable under the control cost approach, as this paper will show.[9]

2 MODEL

This discrete-time model embeds near-rational price adjustment in an otherwise standard New Keynesian general equilibrium framework based on Golosov and Lucas (2007). Retail prices are updated intermittently by a continuum of monopolistically competitive firms. There is also a representative household, and a monetary authority that sets an exogenous growth process for the nominal money supply. All agents act under full information; what this means in our near-rational context will be discussed further below.

2.1 Household

The household's period utility function is $\frac{1}{1-\gamma}C_t^{1-\gamma} - \chi N_t + \nu \ln(M_t/P_t)$, 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

$$\int_0^1 P_{it} C_{it} di + M_t + R_t^{-1} B_t = W_t N_t + M_{t-1} + B_{t-1} + T_t^M + T_t^D, \tag{2}$$

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^M is a lump sum transfer from the central bank, and T_t^D 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

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

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

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

The household's first-order conditions for labor supply, consumption, and money use are:

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

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

$$1 - \frac{\nu P_t}{M_t C_t^{-\gamma}} = \beta E_t \left(\frac{P_t C_{t+1}^{-\gamma}}{P_{t+1} C_t^{-\gamma}} \right). \tag{7}$$

2.2 Monopolistic firms

Each firm i maximizes profits by setting, and (intermittently) resetting, the nominal price P_{it} of its output Y_{it}. It operates the constant returns technology $Y_{it} = A_{it} N_{it}$, where N_{it} is labor, and $A_{it} \equiv \exp(a_{it})$ is an idiosyncratic productivity process. Log productivity a_{it} follows a time-invariant Markov process on a bounded set, $a_{it} \in \Gamma^a \subseteq [\underline{a}, \overline{a}]$, with *i.i.d.* innovations across firms, so a_{it} is correlated with $a_{i,t-1}$, but is uncorrelated with other firms' shocks. Firm i is a monopolistic competitor that faces the demand curve $Y_{it} = C_t P_t^{\epsilon} P_{it}^{-\epsilon}$. Each firm i is infinitesimal, so it assumes that its own price P_{it} has no effect on the aggregate price level P_t. It hires in a competitive labor market at wage rate W_t; its nominal profits per period are:

$$U_{it} = P_{it} Y_{it} - W_t N_{it} = \left(P_{it} - \frac{W_t}{A_{it}} \right) C_t P_t^{\epsilon} P_{it}^{-\epsilon} \equiv P_t u_t(p_{it}, a_{it}). \tag{8}$$

In this equation, $p_{it} \equiv \ln(P_{it}/P_t)$ represents the firm's log real price; the notation $u_t(p_{it}, a_{it})$ indicates real profits per period, as a function of log real price and log productivity.[10,11] Firms are owned by the household, so they discount nominal income between times t and $t+1$ by the factor $\beta \frac{P_t C_{t+1}^{-\gamma}}{P_{t+1} C_t^{-\gamma}}$, consistent with the household's marginal rate of substitution.

Having described the firm's technology and objective function, we next describe its control variable. We assume that the firm's only control is its nominal price P_{it}.[12] It may change this control at any time t, setting a new nominal price $P_{it} \in \Gamma_t^P$, where Γ_t^P is a bounded set further described below. Alternatively, if it takes no action at t, its nominal price remains unchanged, at $\tilde{P}_{it} \equiv P_{i,t-1}$ (tildes indicate prices at the beginning of any period t; end-of-period prices, after any adjustments have occurred, are denoted without tildes). The fact that the firm may deliberately choose a new price $P_{it} \in \Gamma_t^P$, or simply leave it unchanged at $P_{it} = \tilde{P}_{it}$, effectively expands its alternatives to the augmented choice set $\Gamma_t^P(\tilde{P}_{it})$, defined as

$$\Gamma_t^P(\tilde{P}_{it}) \equiv \Gamma_t^P \cup \{\tilde{P}_{it}\} \equiv \Gamma_t^P \cup \{P_{i,t-1}\}. \tag{9}$$

Since the scalar P_{it} is the firm's only control, the quantity it sells is determined by demand: given its price, the firm must fulfill all resulting demand, resulting in gross profits (8).

2.2.1 Price adjustment when choice is costly

We abstract from any cost associated with the act of resetting the price, *per se*.[13] Instead, we assume that *precise decisions are costly*, adopting the "control cost" approach from game theory (see van Damme, 1991, Chapter 4). A key feature of this approach is that decisions are modeled indirectly, "as if" the decision-maker were *selecting a probability distribution* over the feasible set, instead of choosing one of its elements. The decision problem incorporates a cost function that increases with precision: concentrating greater probability on a smaller subset of possible choices increases costs.[14] While control cost models have typically analyzed costly decisions taken at an exogenously fixed point in time, we will also incorporate a cost of *choosing when to make a decision*.[15]

To operationalize this approach, we must define the cost function, which also requires us to choose a definition of precision. Since the time of managers or other employees is a major input to decisions, we define costs in units of time. We define precision in terms of relative entropy, also known as Kullback-Leibler divergence, which measures the deviation between one probability distribution and another. If π_1 and π_2 are the cumulative distribution functions of two different probability distributions defined on the same set, then the Kullback-Leibler divergence $\mathcal{D}(\pi_1||\pi_2)$ of π_1 relative to π_2 is[16]

$$\mathcal{D}(\pi_1||\pi_2) \equiv \int \ln\left(\frac{\pi_1'(x)}{\pi_2'(x)}\right) d\pi_1(x). \tag{10}$$

Following Stahl (1990) and Mattsson and Weibull (2002), we assume that the decision cost is proportional to the Kullback-Leibler divergence of the chosen probability distribution, relative to a default distribution:

ASSUMPTION 1. *The firm's time cost of allocating the probability distribution $\Pi^\dagger$ over the choice set $\Gamma_t^P(\tilde{P})$ is $\kappa\mathcal{D}(\Pi^\dagger||\Theta_t^\dagger(\cdot|\tilde{P}))$, where $\kappa \geq 0$ is a constant, and $\Theta_t^\dagger(\cdot|\tilde{P})$ is a default distribution that conditions on the previous price $\tilde{P}$.*

Assumption 1 implies that there is a default decision $\Pi^\dagger = \Theta_t^\dagger(\cdot|\tilde{P})$ that has zero cost. In order to apply any other probability distribution over the choice set, the firm must pay a positive cost, which we interpret as time devoted to cognitive effort.[17] Here the factor κ represents the marginal cost of entropy reduction, in units of labor time; a higher κ implies that it is more costly to choose the price accurately.

Obviously our model's behavior depends on our assumptions about the default probabilities $\Theta_t^\dagger$ that apply in the absence of cognitive effort. We will impose two key properties on $\Theta_t^\dagger$. First, we assume default behavior is highly random, consistent with the idea that making decisions more precise requires cognitive effort. Second, we impose long-run monetary neutrality, so that our model can be analyzed in real terms after removing a nominal trend. Indeed, even though firms set prices in nominal terms, it is convenient to jump directly to a real description of the firm's problem.[18] So rather than defining decision costs in terms of distributions $\Pi^\dagger$ and $\Theta^\dagger$ over nominal prices, we work with distributions $\pi^\dagger$ and $\theta^\dagger$ that assign the same probabilities to the corresponding sets of real prices.[19] Relative entropy is unchanged by this transformation: $\mathcal{D}(\pi^\dagger||\theta^\dagger) = \mathcal{D}(\Pi^\dagger||\Theta^\dagger)$. Given these preliminaries, we can now define the default distribution, in real terms, conditional on the beginning-of-period real price $\tilde{p}$ (again, tildes distinguish prices at the start of the period).

ASSUMPTION 2. *(a) When adjusting, the firm considers a fixed, bounded set* Γ^p *of log real prices.*

(b) Conditional on adjustment, the firm's default distribution $\theta^\dagger(\cdot|\tilde{p})$ *assigns a uniform distribution* θ *over* Γ^p*.*

(c) The firm's default distribution $\theta^\dagger(\cdot|\tilde{p})$ *associates a constant probability* θ_0 *to nonadjustment.*

Thus, whenever it updates its price, the firm chooses from the same fixed, exogenous set of real prices Γ^p, which we will define wide enough that the real prices preferred at the extremal values of productivity, $\underline{a}$ and $\overline{a}$, lie strictly inside Γ^p.[20] But rather than adjusting, the firm may instead leave its nominal price unchanged, implying the real price $\tilde{p}_{it} \equiv \ln(\tilde{P}_{it}/P_t)$. Thus, as in (9), the firm effectively chooses over the augmented real choice set $\Gamma^p(\tilde{p}_{it})$, defined as

$$\Gamma^p(\tilde{p}_{it}) \;\equiv\; \Gamma^p \cup \{\tilde{p}_{it}\} \equiv\; \Gamma^p \cup \{\ln(P_{i,t-1}/P_t)\}. \tag{11}$$

Together, Assumptions 2(b) and 2(c) imply that the default distribution can be written as

$$\theta^\dagger(p|\tilde{p}) \;=\; (1-\theta_0)\theta(p) + \theta_0 \mathbf{1}(\tilde{p} \leq p), \tag{12}$$

where θ is the *c.d.f.* of a uniform distribution on Γ^p. The assumption of a uniform default distribution is not crucial; what matters qualitatively is just that large, random errors occur if the firm spends no time on its price decision.[21]

FIGURE 1 ABOUT HERE

To describe the firm's behavior under this cost structure, it helps to distinguish the firm's value function at the beginning of t when it still has the option to adjust, $o(\tilde{p}_{it}, a_{it})$, from its value

$v_t(p_{it}, a_{it})$ when production occurs at the end of t, after decisions are made (see the timeline). Following the control cost paradigm, we write the firm's problem "as if" it were choosing a probability distribution over prices, instead of simply choosing a price. It maximizes its expected value, net of computational costs, which are multiplied by the real wage w_t to express all terms in consumption units. Thus, for any bounded choice set Γ, let $\Delta(\Gamma)$ be the set of increasing functions f satisfying $f(\min \Gamma) \geq 0$. Given the beginning-of-t real price $\tilde{p}$, the firm chooses a distribution from $\Delta(\Gamma^p(\tilde{p}))$:

$$o_t(\tilde{p}, a) = \max_{\pi^\dagger \in \Delta(\Gamma^p(\tilde{p}))} \int v_t(p, a) d\pi^\dagger(p) - \kappa w_t \mathcal{D}\left(\pi^\dagger || \theta^\dagger(\cdot|\tilde{p})\right) \quad \text{s.t.} \quad \int d\pi^\dagger(p) = 1, \tag{13}$$

$$v_t(p, a) = u_t(p, a) + E_t \left\{ \beta \frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} o_{t+1}\left(p - i_{t,t+1}, a'\right) \middle| a \right\}. \tag{14}$$

The constraint in (13) states that the chosen function must be a *c.d.f.* (it must integrate to one). Note that since the firm is a nominal price setter, its log real price decreases by the inflation rate $i_{t,t+1} \equiv \ln\left(P_{t+1}/P_t\right)$ between t and $t + 1$.[22]

The decision costs $\kappa w_t \mathcal{D}(\pi^\dagger || \theta^\dagger)$ are a convex function of the chosen probabilities $\pi^\dagger$, the expectation $\int v_t(p, a) d\pi^\dagger(p)$ is a linear function of $\pi^\dagger$, and the simplex on which the probabilities integrate to one is a convex set.[23] Hence (13) maximizes a concave function over a convex set, which implies that a unique solution exists at time t, given the time $t + 1$ value function $o_{t+1}(\tilde{p}_{t+1}, a_{t+1})$.[24] To characterize the probabilities that solve (13), it suffices to define a Lagrangian and take first-order conditions point-by-point at each possible time-t price p. At any p where the default probability $\theta^\dagger$ is differentiable, the distribution $\pi^\dagger$ must satisfy:[25]

$$v_t(p, a) - \kappa w_t \left[\ln\left(\frac{d\pi_t^\dagger(p|\tilde{p}, a)/dp}{d\theta^\dagger(p|\tilde{p})/dp} \right) + 1 \right] - \mu_t(\tilde{p}, a) = 0 \tag{15}$$

where $\mu_t(\tilde{p}, a)$ is the multiplier on the constraint in (13) when the current log real price is $\tilde{p}$, log productivity is a, and the real aggregate state is Ξ_t. Condition (15) implies that the probability density at price p is proportional to an exponential function of $v_t(p,a)$; applying the restriction that probabilities must integrate to one across all p, we conclude that they take the form of a weighted logit. Thus, the probability that the firm selects a real price less than or equal to p is

$$\pi_t^{\dagger}(p|a,\tilde{p}) \equiv \frac{\int^{p} \exp\left(\frac{v_t(p',a)}{\kappa w_t}\right) d\theta^{\dagger}(p'|\tilde{p})}{\int_{\Gamma^p(\tilde{p})} \exp\left(\frac{v_t(p',a)}{\kappa w_t}\right) d\theta^{\dagger}(p'|\tilde{p})}. \tag{16}$$

Moreover, exponentiating (15) and plugging it into the objective function of (13), we obtain an analytical formula for the value function:

$$o_t(\tilde{p},a) = \kappa w_t \ln\left(\int_{\Gamma^p(\tilde{p})} \exp\left(\frac{v_t(p,a)}{\kappa w_t}\right) d\theta^{\dagger}(p|\tilde{p})\right). \tag{17}$$

This solution gives the value of deciding whether or not to adjust the current price, net of decision costs.

2.2.2 Decomposing the timing and pricing decisions

While we have defined the firm's problem as a single decision over the augmented choice set $\Gamma^p(\tilde{p})$, it decomposes naturally into two separate decisions, related to timing and pricing. First, given the beginning-of-period real price $\tilde{p}$, should the firm adjust its price? Second, if it chooses to adjust, what new price should it set? Each of these decisions can be described separately as a costly, error-prone choice, and the resulting policy functions are easier to interpret than (16).

First, consider the timing decision. The value of leaving the beginning-of-t price unchanged is $v_t(\tilde{p},a)$; let $\tilde{v}_t(a)$ represent the value of choosing a new price, conditional on current productivity a. Then, following the control cost approach, we can consider the binary decision to adjust the price (with probability λ_t) or not adjust (with probability $1-\lambda_t$) subject to the time costs $\kappa_\lambda \mathcal{D}((\lambda_t, 1-\lambda_t)||(\bar{\lambda}, 1-\bar{\lambda}))$, where we have defined $\bar{\lambda} \equiv 1-\theta_0$ to represent the default probability of adjusting the price. The choice problem can be written as

$$o_t(\tilde{p},a) \;=\; \max_{\lambda_t}\;\lambda_t\tilde{v}_t(a)+(1-\lambda_t)v_t(\tilde{p},a)-\kappa_\lambda w_t\left[\lambda_t\ln\left(\frac{\lambda_t}{\bar{\lambda}}\right)+(1-\lambda_t)\ln\left(\frac{1-\lambda_t}{1-\bar{\lambda}}\right)\right]. \tag{18}$$

Like (13), problem (18) has a concave objective, since it subtracts convex costs from a function that is otherwise linear in λ_t, so the first-order condition suffices for a maximum of (18). As before, we obtain a weighted logit:[26]

$$\lambda_t(\tilde{p},a) \;\equiv\; \frac{\bar{\lambda}\exp\left(\frac{\tilde{v}_t(a)}{\kappa_\lambda w_t}\right)}{\bar{\lambda}\exp\left(\frac{\tilde{v}_t(a)}{\kappa_\lambda w_t}\right)+(1-\bar{\lambda})\exp\left(\frac{v_t(\tilde{p},a)}{\kappa_\lambda w_t}\right)} \;\in\; [0,1]. \tag{19}$$

The weight parameter $\bar{\lambda}$ is related to the speed of choice. In particular, it represents the update probability in one discrete time period when the firm is indifferent between adjusting or not: $\lambda_t(\tilde{p},a)=\bar{\lambda}$ if $v_t(\tilde{p},a)=\tilde{v}_t(a)$.

Next, consider the pricing decision. If the firm decides to adjust, it chooses a price from the set Γ^p, subject to decision costs, as follows:

$$\tilde{v}_t(a)=\max_{\pi\in\Delta(\Gamma^p)}\int v_t(p,a)d\pi(p)-\kappa_\pi w_t\mathcal{D}\left(\pi||\theta\right) \quad \text{s.t.} \quad \int d\pi(p)=1\,. \tag{20}$$

Again, as in (13), we are maximizing a concave function on a convex set. Taking first-order conditions, we find a weighted logit analogous to (16):[27]

$$\pi_t(p|a) \;\equiv\; \frac{\int^p \exp\left(\frac{v_t(p',a)}{\kappa_\pi w_t}\right)d\theta(p')}{\int_{\Gamma^p}\exp\left(\frac{v_t(p',a)}{\kappa_\pi w_t}\right)d\theta(p')} \tag{21}$$

The parameter κ_π in the logit function can be interpreted as the degree of noise in the price decision; in the limit as $\kappa_\pi\to 0$, (21) converges to the policy function under full rationality, so that the optimal price is chosen with probability one.

In the particular case $\kappa_\lambda = \kappa_\pi \equiv \kappa$, the policy functions (19) and (21) are equivalent to (16). Concretely, wherever the densities of π_t and $\pi_t^\dagger$ exist, they satisfy $\frac{d\pi_t^\dagger(p|\tilde{p},a)}{dp} = \lambda_t(\tilde{p},a)\frac{d\pi_t(\tilde{p},a)}{dp}$. The probability that a firm leaves its nominal price unchanged is given by $1-\lambda_t(\tilde{p},a)$, corresponding to a point mass in the distribution $\pi_t^\dagger$ at the unadjusted price $\tilde{p}$.

But note that in (18)-(21) we generalize our setup to differentiate between the noise parameters in the timing (κ_λ) and pricing (κ_π) decisions. This generalization allows us to study costs and errors in the pricing decision separately from costs and errors in the timing decision. Going forward, we will compare a version of our model with "errors in pricing" (EiP) only, setting $\kappa_\pi > 0$ but $\kappa_\lambda = 0$, and a version with "errors in timing" (EiT) only, setting $\kappa_\pi = 0$ but $\kappa_\lambda > 0$. These two frictions have very different effects, but both, jointly, help explain price dynamics.

2.2.3 Discussion

Some interpretive comments are useful here. First, although we write the decision problem "as if" the firm chooses a probability distribution over prices, this should not be taken literally. A more realistic interpretation is that the firm chooses how much time to dedicate to its decision, and thereby achieves more or less precision in its price choice. Thus, defining the problem as a choice of a mixed strategy is just a way to incorporate errors into the model. And writing it as an optimization problem just serves to discipline the errors, implying that more costly mistakes are less likely. Features of the model that we do take seriously include (a) choice is costly in terms of time and other resources; (b) therefore decision-makers sometimes fail to take the action that would otherwise be optimal; (c) *ceteris paribus*, more valuable actions are more likely than less valuable ones; (d) in a retail pricing context, errors affect both the timing of price adjustment, and the actual price chosen; and therefore (e) firms will try to balance the marginal benefits and costs of precision across these two margins. Quantitatively, we will see that this framework yields a remarkably successful model of nominal rigidity, even though we impose strong restrictions on the form of the control cost function.

Second, we reiterate that "rational inattention" and the control costs framework address different "stages" of the choice process: obtaining and using information.[28] Rational inattention imposes a cost function that depends on the amount of information processed while making decisions. Instead, we assume that the firm has enough information so that choosing the optimal action is feasible, but that actually inferring the best action from its information is costly. Thus, while choices in a rational inattention model are conditioned on a prior, here problem (13) conditions instead on the firm's true state variables (its lagged price $\tilde{p}$, its productivity a, and the aggregate state Ξ). So when the true value $v_t(p, a)$ of each possible price p appears in (13), this does not mean the firm "knows" the true values, but only that it has enough information to make calculating $v_t(p, a)$ feasible.[29] By devoting more time to its decision, the firm could take more factors into consideration (incorporating additional variables into its calculations), or study those factors more carefully (checking its calculations, or considering higher-order terms). But given our assumed cost function, calculating $v_t(p, a)$ to full precision is not worthwhile, so the firm prefers to conserve managerial time by tolerating some imprecision in its choices.

2.3 Distributional dynamics

As firms respond to productivity shocks, managing their prices according to (19) and (21), the distribution of prices and productivities evolves over time. Recall that $\tilde{p}_{it}$ refers to firm i's log real price at the beginning of t, prior to adjustment; this may of course differ from the log real price p_{it} at which it produces, since it may reset its nominal price before production. Therefore we will distinguish the beginning-of-t distribution of prices and productivity, $\widetilde{\Psi}_t(\tilde{p}_{it}, a_{it})$, from the distribution at the time of production, $\Psi_t(p_{it}, a_{it})$.

Two stochastic processes drive the distributional dynamics. First, there is the Markov process for firm-specific productivity, which we can write in terms of the following *c.d.f.*:

$$S(a'|a) \;=\; prob(a_{i,t+1} \leq a'|a_{it} = a). \tag{22}$$

Thus, suppose the *c.d.f.* of log real prices and productivities at the end of $t-1$ is $\Psi_{t-1}(p,a)$. This distribution is then hit by the productivity process (22). Moreover, fixing a firm's nominal price, its log real price is shifted by inflation, from $p_{i,t-1}$ to $\tilde{p}_{it} \equiv p_{i,t-1} - i_{t-1,t}$, where $i_{t-1,t} \equiv \ln(P_t/P_{t-1})$. Thus the distribution of log real prices and productivities at the start of t is:[30]

$$\widetilde{\Psi}_t\left(\tilde{p},a'\right) \;=\; \int S(a'|a)\, d_a \Psi_{t-1}(\tilde{p}+i_{t-1,t},a). \tag{23}$$

The second stochastic process that drives the distributions is the price updating process. A firm with log real price $\tilde{p}$ and log productivity a at the start of t adjusts its price with probability $\lambda_t(\tilde{p},a)$, given by (19), and upon adjustment its new log real price is conditionally distributed according to $\pi_t(p|a)$, given by (21). Therefore, if the beginning-of-t distribution of firm states is $\widetilde{\Psi}_t(\tilde{p},a)$, the distribution at the end of t is

$$\Psi_t(p,a) \;=\; \int^p \left(1-\lambda_t(\tilde{p},a)\right) d_{\tilde{p}}\widetilde{\Psi}_t(\tilde{p},a) + \int \lambda_t(\tilde{p},a)\pi_t(p|a)\, d_{\tilde{p}}\widetilde{\Psi}_t(\tilde{p},a). \tag{24}$$

2.4 Monetary policy and general equilibrium

The nominal money supply is shocked by an AR(1) process z,[31]

$$z_t = \phi_z z_{t-1} + \epsilon^z_t, \tag{25}$$

where $0 \le \phi_z < 1$ and $\epsilon^z_t \sim i.i.d.N(0,\sigma_z^2)$. Here z_t is the time t rate of money growth:

$$M_t/M_{t-1} \equiv \mu_t = \mu \exp(z_t). \tag{26}$$

Seigniorage revenues are paid to the household as a lump sum transfer T_t^M, and the government budget is balanced each period, so that $M_t = M_{t-1} + T_t^M$. Bond market clearing implies $B_t = 0$.

We have now discussed all the ingredients of general equilibrium. We will define equilibrium in real terms, without reference to the nominal price level P_t, which is nonstationary under the money supply rule (26). To do so, we must define a real aggregate state Ξ_t that summarizes all predetermined variables needed to calculate equilibrium objects at t. Note that a firm's real beginning-of-period price $\tilde{p}_{it} \equiv \ln(P_{it}/P_t)$ is *not* predetermined at t, since it depends on the aggregate price level P_t, which is endogenous at t. Thus the beginning-of-t real distribution $\widetilde{\Psi}_t$ is likewise not predetermined at t, and cannot form part of of the time-t real state variable Ξ_t. Instead, we will construct an equilibrium in terms of the real state $\Xi_t \equiv (z_t, \Psi_{t-1})$, which is predetermined at t since it depends only on the lagged distribution Ψ_{t-1}.

Recall that time subscripts on aggregate variables denote dependence on the aggregate state Ξ_t; for example, we write consumption as $C_t \equiv C(\Xi_t)$, and the end-of-t value function as $v_t(p,a) \equiv v(p,a,\Xi_t)$. Calculating general equilibrium requires us to find value functions v_t, o_t, and $\tilde{v}_t$ that satisfy (14), (18), and (20). These value functions are associated with policy functions λ_t and π_t that satisfy (19) and (21). These policy functions drive the dynamics of the distributions $\widetilde{\Psi}_t$ and Ψ_t according to (23) and (24). Note in particular that if the functions λ_t and π_t are known, then (23) and (24) can be used to calculate Ψ_t from Ψ_{t-1}, which is the key to updating the state Ξ_t.

Besides these functions, general equilibrium also involves several scalar processes that must obey the household's first-order conditions. Given money growth (26), real money demand $m_t \equiv M_t/P_t$, the real wage $w_t \equiv W_t/P_t$, and consumption C_t must satisfy:

$$\mu \exp(z_t - i_{t-1,t}) = \frac{m_t}{m_{t-1}}, \tag{27}$$

$$w_t C_t^{-\gamma} = \chi, \tag{28}$$

$$1 - \frac{\nu}{m_t C_t^{-\gamma}} = \beta E_t \left(\frac{C_{t+1}^{-\gamma}}{i_{t,t+1} C_t^{-\gamma}} \right). \tag{29}$$

Although the nominal price level never appears in the real equilibrium equation system, the inflation rate $i_{t-1,t} \equiv \ln(P_t/P_{t-1})$ does. Inflation must satisfy (4), which defines the aggregate price level, and reduces to the following identity in real terms:

$$\int\int \exp((1-\epsilon)p)\, d_p d_a \Psi_t(p,a) \;=\; 1. \tag{30}$$

This gives us enough equations to determine the value functions, policy functions, and distributions, as well as the scalars m_t, w_t, C_t, and $i_{t-1,t}$. A solution to these equations constitutes a real general equilibrium of the economy. This requires a high-dimensional calculation, because the aggregate state Ξ_t includes the distribution of firms' individual state variables. We compute the model using the algorithm of Reiter (2009), as described in Online Appendix A.

Note that computing general equilibrium does not require us to calculate labor, because of our linear disutility assumption. Labor supply is perfectly elastic, adjusting for consistency with consumption and the real wage, so it can be calculated from goods market clearing:

$$N_t - K_t^\lambda - K_t^\pi = \int_0^1 \frac{C_{it}}{A_{it}} di = C_t \int\int \exp(-\epsilon p - a)\, d_{\tilde{p}} d_a \Psi_t(p,a) \equiv \Delta_t C_t, \tag{31}$$

where K_t^λ is total time devoted to deciding whether to adjust prices, and K_t^π is total time devoted to choosing which price to set by firms that adjust. These quantities are given by

$$\begin{aligned} K_t^\pi &\equiv \kappa_\pi \int\int \lambda_t(\tilde{p},a) \left(\int_{p\in\Gamma^p} \ln \pi(p) d_p \pi(p) - \ln \bar{u} \right) d_{\tilde{p}} d_a \tilde{\Psi}_t(\tilde{p},a) \\ K_t^\lambda &\equiv \kappa_\lambda \int\int \left(\lambda_t(\tilde{p},a) \ln \frac{\lambda_t(\tilde{p},a)}{\bar{\lambda}} + (1-\lambda_t(\tilde{p},a)) \ln \frac{1-\lambda_t(\tilde{p},a)}{1-\bar{\lambda}} \right) d_{\tilde{p}} d_a \tilde{\Psi}_t(\tilde{p},a) \end{aligned}$$

Equation (31) also defines a measure of price dispersion,

$$\Delta_t \;\equiv\; \int\int \exp(-\epsilon p - a)\, d_p d_a \Psi_t(p,a) \tag{32}$$

weighted to allow for heterogeneous productivity. As in Yun (2005), an increase in Δ_t decreases the goods produced per unit of labor, rather like a negative aggregate productivity shock.

3 RESULTS

We now describe the model calibration and report simulation results, documenting steady state micro implications for price adjustments, and macro implications for monetary policy shocks and changes in trend inflation. The simulations are performed at monthly frequency, and all data and model statistics are monthly unless stated otherwise.

To better understand our model's behavior, we will also decompose it to look separately at the two margins of error: mistakes in prices, and mistakes in timing. We report a special case that allows for errors in the size of price changes, but not in their timing, imposing $\kappa_\lambda = 0$ but allowing $\kappa_\pi > 0$; this is labelled "errors-in-prices" (EiP).[32] In the EiP version, a firm will adjust its nominal price at time t if and only if the value of adjustment, $\tilde{v}_t(a)$, exceeds the value $v_t(\tilde{p}, a)$ of maintaining its current price $\tilde{p}$. At the opposite extreme, we also consider a case with errors in the timing of price adjustments, but not in their size, imposing $\kappa_\pi = 0$ but allowing $\kappa_\lambda > 0$; this case is labelled "errors-in-timing" (EiT). The version that nests both types of errors, which imposes $\kappa_\pi = \kappa_\lambda \equiv \kappa > 0$, is labelled "nested".

Likewise, for the sake of comparison, we will also report a Calvo (1983) specification and a fixed menu cost (FMC) model. Whenever we refer to the "main model" or the "benchmark model", we mean the nested specification.

3.1 Parameters

The key parameters of the decision process are the rate and noise parameters $\bar{\lambda}$ and κ. We estimate these parameters to match two steady-state features of retail pricing data: the average rate of adjustment, and the histogram of nonzero log price adjustments. For the estimates we use the Dominick's supermarket dataset described in Midrigan (2011), after removal of price changes related to "sales", and aggregating weekly adjustment rates to monthly rates for comparability with some of the other data sources we consider. Our reason for ignoring sales is that recent literature has found that monetary nonneutrality depends primarily on the frequency of "regular" or "non-sale" price changes (see for example Eichenbaum *et al.* 2011; Guimaraes and Sheedy 2011; or Kehoe and Midrigan 2015).

More precisely, let h be a vector of length $\#h$ representing the probabilities of nonzero log price adjustments in a histogram with $\#h$ fixed bins.[33] We choose the adjustment parameters $\bar{\lambda}$ and κ (or κ only in the EiP version) to minimize the following distance criterion:

$$\text{distance} \;=\; \sqrt{\#h}\,||\lambda_{model} - \lambda_{data}|| + ||h_{model} - h_{data}|| \tag{33}$$

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 data, and h_{model} and h_{data} are the vectors of bin probabilities for nonzero price adjustments in the model and the data.[34] The nonzero adjustments in the histogram h_{data} are "regular" price changes from the Dominick's dataset. The adjustment frequency $\lambda_{data} = 10.2\%$ represents the unweighted mean weekly adjustment hazard in the same data, rescaled to monthly units. Clearly these features of the data are informative about the two parameters, since $\bar{\lambda}$ will shift the adjustment hazard and κ will spread out the distribution of price changes.

The rest of the parameterization is less crucial for our purposes, so we adopt utility parameters from Golosov and Lucas (2007), including the discount factor, $\beta = 1.04^{-1/12}$; the coefficient of relative risk aversion on consumption, $\gamma = 2$; and the elasticity of substitution in the consumption aggregator, $\epsilon = 7$. Likewise, we set the marginal disutility of labor to $\chi = 6$, and the coefficient on the utility of money is $\nu = 1$.

We assume productivity is AR(1) in logs:

$$\ln A_{it} = \rho \ln A_{it-1} + \varepsilon_t^a, \tag{34}$$

where ε_t^a is a mean-zero, normal, *iid* shock. For numerical purposes, (34) is approximated by a finite Markov chain, following the Tauchen method. We set ρ following Blundell and Bond (2000), who estimate an annual autocorrelation of 0.565 in a panel of US manufacturing companies from 1982 to 1989; this implies approximately $\rho = 0.95$ at monthly frequency. The variance of log productivity is $\sigma_a^2 = (1-\rho^2)^{-1}\sigma_\varepsilon^2$, where σ_ε^2 is the variance of the innovation ε_t^a. We set $\sigma_a = 0.06$, which is the standard deviation of "reference costs" estimated by Eichenbaum *et al.* (2011). Our Markov chain allows for four standard deviations of productivity, so in the simulations $a \equiv \ln A$ lies in $[\underline{a}, \overline{a}] \equiv [-0.24, 0.24]$. We compare productivity to "reference costs" rather than "weekly costs" because this is more closely comparable to the pricing data we consider, which exclude sales. The rate of money growth is set to match the 2% annual inflation rate observed in the Dominick's dataset.

The width of the uniform default distribution θ in the decision cost function (see Assumption 2) is also a free parameter of our setup, but it is not a crucial one. We assume the support of θ stretches 25% beyond the range of real prices that would be chosen if pricing were perfectly frictionless; that is, θ is uniform on $\left[\ln\left(\frac{\epsilon w}{\epsilon-1}\right) - 1.25\overline{a}, \ln\left(\frac{\epsilon w}{\epsilon-1}\right) - 1.25\underline{a}\right]$. Online Appendix C shows that our results are robust to changing the support of θ.[35] We have also run robustness exercises that replace the uniform default θ with a truncated normal having the same standard deviation, but there is little change in the results. Also, like the cost shock, real prices in our simulations are represented on a finite grid. In our reported simulations the step sizes in the price and productivity grids are the same; since actual price changes are typically much larger than this, we find that making the price grid finer has a negligible effect on the results.

TABLE 1 ABOUT HERE

Parameter estimates for all the specifications we compare are reported in Table 1.[36] The EiP specification has only one free parameter: the pricing noise κ_π. The errors-in-timing model has two free parameters: the rate parameter $\bar{\lambda}$, and the timing noise κ_λ. The nested model features the same two free parameters, but the noise parameter applies both to the timing and pricing decisions ($\kappa_\pi = \kappa_\lambda \equiv \kappa$).[37] Overall we will see (Table 2) that our estimates imply a low noise level, causing only modest revenue losses. The rate parameter $\bar{\lambda}$ is estimated to be less than the observed adjustment frequency in the errors-in-timing specification, but is twice as high as the observed adjustment frequency in the main nested model. The combination of a high underlying adjustment rate, together with a low noise parameter, indicates that our benchmark model is consistent with a high degree of rationality.

3.2 Policy functions

The steady state policy functions of the main model are illustrated in Figure 2. The first panel illustrates the distribution of reset prices 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 any given a reflects the presence of errors. If we instead eliminate pricing errors by considering the limit $\kappa_\pi \to 0$ (the EiT model), then the bell-shaped price distribution collapses to a sharp peak that places all the probability mass on the conditionally-optimal price.

Similarly, the top right panel shows the probability $\lambda(\tilde{p}, a)$ of price adjustment conditional on beginning-of-period price and productivity, which looks rather like a vertical reflection of the probability surface $\pi(p|a)$. Near (but not exactly on) the 45^o-line, the adjustment probability reaches a (strictly positive) minimum; moving away from this minimum, it increases smoothly towards one. Thus, the prices p that the firm is most likely to choose, conditional on log productivity a, are also those that it is least likely to alter, conditional on the same a. Again, it is instructive to consider how the policy function changes if we eliminate errors. If we eliminate timing errors by considering the limit $\kappa_\lambda \to 0$ (the EiP model), then the smoothly state-dependent adjustment hazard collapses to a pair of (S,s) bands. That is, when $\kappa_\lambda = 0$, the firm adjusts its price with probability one if and only if the current price lies sufficiently far from the conditional optimum; otherwise it adjusts with probability zero.[38]

FIGURE 2 ABOUT HERE

The bottom panels of Figure 2 plot the same policies in greater detail, by slicing through the three-dimensional policy functions at each point a on the grid of possible log productivities; the lowest productivity ($\underline{a} = -0.24$) is highlighted with stars, while the highest productivity ($\overline{a} = 0.24$) is highlighted with squares. For ease of comparison, each slice is centered around the 45^o line of the upper panels, so the variable on the horizontal axis is $p - \ln\left(\frac{\epsilon w}{\epsilon - 1}\right) + a$, which is the price as a log deviation from its flexible-price optimum. These two-dimensional slices make it easier to see the size of the errors and the strength of the selection effect in equilibrium. The bottom left panel shows that at the minimum cost level (squares), prices vary by roughly $\pm 5\%$ around their optimal value. At the highest costs (stars), errors are on the order of $\pm 10\%$. This reflects the fact that high costs imply a high optimal price, so the firm sells at low volume; hence pricing errors have a smaller impact on profits, so the firm tolerates larger errors. We also see that the mode of the low (high) cost curve lies to the right (left) of zero; in other words, the preferred price is not exactly on the 45^o line of the upper panels. This occurs because a firm with nominal rigidities and autoregressive shocks must set prices "conservatively", responding slightly less than proportionally to costs. When costs are unusually low (high), the firm expects them to rise (fall), so it prefers a higher (lower) price than it would if it expected to adjust every period.

Likewise, the lower right panel illustrates the selection effect. We see that with low costs (squares), the monthly hazard rate reaches a minimum of 3% when the price is 4% above the 45^o line. The hazard rate rises steeply away from this minimum; indeed, selection is strong here compared with the evidence of Campbell and Eden (2014, Figure 3) or Eichenbaum *et al.* (2011, Figs. 8-10). In contrast, with high costs (stars) the hazard rate is flatter overall, reflecting a weaker selection effect when the volume of sales is lower. In other words, just as prices are optimally less accurate when costs are high, the selection effect is also weaker (timing is less accurate) when costs are high, because the profits at stake are lower. The asymmetry in the hazard function (and the price distribution) when costs are high (stars) is a related effect. An excessively high price, leading to low sales, is a less costly error than an excessively low price (implying high sales at a cheap price). Therefore the hazard function is steeper on the left, as prices fall below their target, than it is on the right, as prices exceed their target.

3.3 Distribution of price adjustments

Table 2 documents the steady state distributions of price adjustments in the EiT, EiP, and nested specifications defined above, as well as the FMC and Calvo models, and compares them to the Dominick's data. Figure 3 makes the same comparison graphically, showing the histograms of nonzero log price changes from the model simulations and the corresponding histogram from the data. The vector of bin frequencies for the 81 bars in these histograms is the object that enters the second term of the distance criterion (33) used to estimate our models.

TABLE 2 ABOUT HERE

The results demonstrate the potential of timing and pricing errors to spread out the distribution of price changes and thereby better match the empirical histogram. All five model versions match the 10.2% monthly adjustment frequency observed in the data. But with only one free parameter, the EiP model then lacks an additional degree of freedom to fit the typical size of price changes. The EiP estimate has very low noise ($\kappa_\pi = 0.008$), resulting in behavior that is quite close to full rationality. The implied distribution of price changes thus resembles the FMC case, with two sharp spikes representing increases or decreases occurring near a pair of (S,s) bands. Pricing errors spread out the spikes slightly in the EiP model, so there is some variation in the size of price increases, and in the size of price decreases. But this variation is less than we observe in the Dominick's data; price changes in the EiP and FMC models are too small on average, and there is little mass in the tails of the distribution.

Since the errors-in-timing model has two free parameters, it might be expected to fit both the frequency and size of adjustments. But neither $\bar{\lambda}$ nor κ_λ acts to spread out the distribution in a way that fits the data; the tails of the distribution of adjustments drop off very steeply in the EiT simulation. This contrasts with our Calvo simulation, which shows a tall spike of near-zero price adjustments (71% of changes are smaller than 5% in absolute value), but also has fatter tails than EiT. The difference between these two specifications relates to state dependence in the adjustment hazard. Prices are unlikely to drift far from their target levels in the EiT model, since the adjustment hazard increases sharply as the value of adjustment rises, but the Calvo adjustment hazard is always a constant 10.2% per month, so the price occasionally drifts far from its target, resulting in a large change.

FIGURE 3 ABOUT HERE

In contrast, the coexistence of timing and pricing errors in the nested model helps it fit both the frequency and the size of adjustments in the Dominicks's data. Very small price changes may occur in two main ways: firms may correctly choose a small adjustment when they mistakenly thought an adjustment was urgent; or they may mistakenly calculate that they only require a small change when in fact they urgently need a large one.[39,40] The very largest changes in the distribution are likely to reflect mistaken overcompensation to a failure to adjust to a series of large cost shocks— that is, a pricing error reinforcing a timing error.

Together, these effects generate a wide, fat-tailed distribution that matches both the typical frequency and the typical size of price adjustments. Thus the main model is fairly consistent with the mean absolute change, the standard deviation of the adjustments, the fraction of small changes, and even the kurtosis of the data. (Midrigan, 2011, instead imposes a leptokurtic shock process directly.) Thus, while the restriction $\kappa_\pi = 0$ that defines the EiT specification strongly constrains its ability to match the data, the restriction $\kappa_\pi = \kappa_\lambda$ that we impose on the nested model seems largely consistent with the evidence.

Looking at the fat tails in Figure 3, it might seem that firms in our setup often make large mistakes. But Figure 4 shows that mistakes are rarely costly. It shows the distribution of losses

$d(p, a) \equiv \tilde{v}(a) - v(p, a)$ from nonadjustment under the benchmark specification, expressed as a percentage of average monthly revenue, both at the beginning and at the end of the period. The distribution of losses is strongly skewed out to the right: losses of up to 8% of revenue are visible in the histogram, but most of the mass is concentrated at the left, with a mode at *negative* 7%. In other words, the firms at the left end of this distribution strictly prefer not adjusting, because adjustment would imply paying a decision cost, and would also imply a risk of setting the wrong price (this latter phenomenon is what we call "precautionary price stickiness" in Costain and Nakov, 2015A). Adjustment eliminates some, but not all, of the largest losses, so the beginning-of-period distribution ($\tilde{\phi}(\tilde{p}, a)$, shown as a shaded area) shifts slightly leftward ($\phi(p, a)$, shown as a solid line) before production and transactions occur. Adjustment fails to eliminate the right tail of the distribution for two reasons: some firms that would potentially benefit from adjustment fail to adjust, and some that do adjust make costly errors.

FIGURE 4 ABOUT HERE

Another measure of the losses due to costly choice is reported at the end of Table 2. The last line of the table shows the average monthly gain from eliminating all decision costs and frictions, as a fraction of average monthly revenues.[41] The previous two lines decompose the losses, showing the costs K^π of choosing prices and the costs K^λ of deciding the timing of adjustment. The difference between the total loss in the last line of the table, and the sum $K^\pi + K^\lambda$, represents the cost of errors. The largest total loss occurs in the nested model, where choosing prices costs firms half of one percent of revenues, choosing the timing of adjustment costs one-third of one percent of revenues, and errors eat up another half a percent of revenues.[42] In a case study of an industrial firm, Zbaracki *et al.* (2004) find that decision and negotiation costs associated with price adjustment eat up roughly 1.2% of revenues; this is larger than the decision costs, 0.87%, that we find for the nested model.[43] They do not attempt to calculate the revenue loss caused by the suboptimality of the price process at the firm they study.

3.4 Some puzzles from microdata

Our model also helps explain a number of puzzling observations from microdata. First, note that our main ("nested") model reproduces Eichenbaum, Jaimovich, and Rebelo's (2011) finding that prices are more volatile than costs (see Table 2). In their data, the ratio of the standard deviation of log reference prices to that of log reference costs is 1.15.[44] Both the FMC and Calvo models predict that this ratio should be less than one, because optimal prices in these frameworks anticipate mean reversion of productivity shocks. In our simulations, the ratio is 0.95 in the FMC model and 0.77 in the Calvo model; pricing is more "conservative" in the Calvo case because there is no state dependence in the adjustment hazard. Likewise, prices are less volatile than costs in the EiP and especially the EiT versions of our model. But in the nested version, pricing and timing errors interact to augment price dispersion. In particular, when a delayed adjustment to an exogenous shock (a timing error) is reinforced by a pricing error in the same direction, the result is an exceptionally large adjustment. Under our calibration, the interaction of these two types of errors makes prices more variable than costs, as in the data.

FIGURE 5 ABOUT HERE

Figure 5 shows how our benchmark model performs relative to some microdata facts that condition on the time since last adjustment.[45] First, one might intuitively expect adjustment hazards to increase with the time since the last change. Indeed, this is what happens in the EiT and FMC specifications, in which newly-set prices are conditionally optimal, and subsequent inflation and productivity shocks gradually drive prices out of line with costs. But most empirical studies find that price adjustment hazards are mildly *decreasing* with the time since adjustment, even after controlling for heterogeneity, as in Nakamura and Steinsson (2008), whose empirical hazards are shown as shaded bars in the left panel of Figure 5. The nested specification is more consistent with these data, since it implies that the hazard is largely independent of the age of the price. In fact, the hazard in this version of the model has a mildly negative slope at the beginning, because firms sometimes choose to readjust to correct a recent mistake.

The shaded bars in the middle panel of Figure 5 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 slightly positive slope. In the EiT and FMC models (not shown), 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 errors-in-prices 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 errors). It is unclear which of our specifications performs best relative to this feature of the microdata.

Finally, the right panel of the figure illustrates Campbell and Eden's (2014) observation that extreme prices tend to be young. The shaded 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 errors-in-prices and nested models; in these models extreme prices often result from an extreme productivity draw compounded by an error, and are therefore unlikely to last long.

3.5 Money supply shocks

Next, we turn to the effects of monetary shocks. Figure 6 shows the impulse responses of inflation and consumption to a 1% money growth rate shock with monthly autocorrelation 0.8, in each parameterization of our model and also in the Calvo and FMC frameworks. As Golosov and Lucas (2007) and other recent papers have made clear, the strength of monetary nonneutrality varies greatly across models of price stickiness. After an increase in money growth, the

Calvo model (highlighted with diamonds) implies a small but very persistent rise in inflation, leading to a large and persistent expansion of output. In the FMC model (triangles), there is instead a large inflation spike that is even less persistent than the money growth process itself, as price adjustment anticipates the autocorrelation of money growth; this abrupt equilibration of prices makes the effect on output smaller and much less persistent.

FIGURE 6 ABOUT HERE

The real effects arising in our benchmark nested model are intermediate between the Calvo and FMC cases. Consumption rises by 1.8% on impact in the benchmark model (highlighted with dots) following a one percent money growth shock, and converges back to steady state with a half-life of four months. This is less persistence than we reported for the "smoothly state-dependent pricing" specification of Costain and Nakov (2011B), but it is roughly twice as persistent as the real response of the FMC model. If we take the area under the consumption impulse response function as a measure of the total nonneutrality, then the figures show that our nested model has roughly two-and-a-half times the nonneutrality of the FMC case, while in turn the Calvo framework triples the nonneutrality again.

Interestingly, the impulse responses in the EiT case (squares) closely resemble those of the nested model, with a similarly strong real expansion. This suggests that the timing errors associated with the logit hazard function are the main factor responsible for the real effects in the nested model too. Timing errors obviously help drive monetary nonneutrality since they imply that some prices fail to adjust immediately after a monetary shock. In other words, timing errors reduce the "selection effect" that eliminates the reaction of real variables to nominal shocks in some state-dependent pricing models. The complete lack of a selection effect explains the strength of nonneutrality in the Calvo model.

In contrast, money growth shocks in the EiP framework (circles) have a much weaker impact on consumption. Pricing errors are small, and timing is perfectly rational, so the small decision cost and the risk associated with price adjustment in the EiP model basically act like a small menu cost. Thus, as we already saw in Figure 3, the EiP and FMC specifications behave very similarly when they are calibrated to the same data. This is true of the impulse responses too: a money supply shock causes a strong initial inflation spike, due to the immediate price changes made as firms cross the lower (S,s) band when the money supply increases, and the impulse responses are almost indistinguishable across the two specifications.

FIGURE 7 ABOUT HERE

To show that the brief inflation spike in the EiP and FMC models is indeed due to "selection" in the sense of Golosov and Lucas (2007), we decompose the inflation response in Figure 7. We define the conditionally optimal price $p_t^*(a) \equiv \text{argmax}_p v_t(p,a)$, and also $x_t^*(\tilde{p},a) \equiv p_t^*(a) - \tilde{p}$, the desired log price change of a firm that starts period t with log productivity a and log real price $\tilde{p}$. Firm i's actual adjustment can thus be decomposed as $x_{it} = x_t^*(\tilde{p},a) + \epsilon_{it}$, where ϵ_{it} is an error, in logs. We can write the average desired adjustment $\overline{x}_t^*$, the fraction of firms adjusting $\overline{\lambda}_t$, and the average log error $\overline{\epsilon}_t$ as follows:

$$\overline{x}_t^* = \int\int x_t^*(\tilde{p},a) d_{\tilde{p}} d_a \widetilde{\Psi}_t(\tilde{p},a)\,, \tag{35}$$

$$\overline{\lambda}_t = \int\int \lambda_t(\tilde{p},a) d_{\tilde{p}} d_a \widetilde{\Psi}_t(\tilde{p},a)\,, \tag{36}$$

$$\overline{\epsilon}_t = \int\int \left\{ \int (p - p_t^*(a))\, d_p \pi_t(p|a) \right\} \lambda_t(\tilde{p},a)\, d_{\tilde{p}} d_a \widetilde{\Psi}_t(\tilde{p},a). \tag{37}$$

Then inflation can be expressed as

$$i_t \;=\; \int\int x_t^*(\tilde{p},a)\lambda_t(\tilde{p},a)\, d_{\tilde{p}} d_a \widetilde{\Psi}_t(\tilde{p},a) + \overline{\epsilon}_t. \tag{38}$$

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

$$\Delta i_t = \overline{\lambda}\Delta\overline{x}_t^* + \overline{x}^*\Delta\overline{\lambda}_t + \Delta \int\int x_t^*(\tilde{p},a)\left(\lambda_t(\tilde{p},a) - \overline{\lambda}_t\right) d_{\tilde{p}} d_a \widetilde{\Psi}_t(\tilde{p},a) + \Delta\overline{\epsilon}_t, \tag{39}$$

where terms without time subscripts represent steady states, and Δ represents a change relative to steady state.[46]

TABLE 3 ABOUT HERE

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 population. The "selection effect", $\mathcal{S}_t \equiv \Delta \int\int x^*_t(\tilde{p},a)\left(\lambda_t(\tilde{p},a) - \overline{\lambda}_t\right) d_{\tilde{p}} d_a \widetilde{\Psi}_t(\tilde{p},a)$, is the inflation caused by redistributing adjustment opportunities from firms desiring small (or negative) price changes to firms desiring large (positive) changes, while fixing the number of adjusters. The last term, $\Delta\overline{\epsilon}_t$, is the change in the average log error. Figure 7 decomposes inflation in the EiP, EiT, and nested models. As expected, the spike of inflation on impact in EiP is overwhelmingly attributed to selection. Interestingly, inflation is also driven primarily by selection in the nested model, but this selection effect is more spread out over time. The intensive margin is smaller, and the extensive margin and error margins are negligible, in all cases considered.[47]

In Table 3, we further assess the degree of monetary nonneutrality by running two calculations from Golosov and Lucas (2007). Assuming for simplicity that inflation is driven by money shocks only, we calibrate the variance of the shocks for each specification to fit the standard deviation of quarterly US inflation (0.25%). We then check what fraction of US output fluctuations can be explained by those shocks. In the EiT and nested cases, these money shocks would explain more than 70% or 80% of the observed variation in US output growth, while in EiP they would explain only 38%, consistent with the strong inflation spike and small output response observed in Figure 6 for EiP. In the last line of the table, we also report "Phillips curve" coefficients, meaning estimates from an IV regression of the effect of inflation on output, instrumenting inflation by the exogenous money supply process. The coefficient is twice as large for the nested and errors-in-timing cases as it is for EiP. Thus, allowing for errors in timing suffices to generate large real effects of money shocks.

FIGURE 8 ABOUT HERE

Figure 8 shows that nonneutrality in the nested model is quite robust to large changes in the persistence of money growth. It compares impulse responses as the persistence parameter ϕ_z varies from 0 (unadorned line) to 0.8 monthly (the benchmark value, shown as a line with dots), and to 0.9, 0.99, and 0.999 monthly (triangles, squares, and circles, respectively). In each case the amplitude of the shock is rescaled to give a five percent total increase in the money supply. To emphasize the scale of the shock, we plot all the impulse responses in cumulative terms, showing the level of money, the level of prices, and the cumulative change in consumption.[48] As monetary persistence increases, the real effects of the shock gradually vanish. Qualitatively, this is unsurprising; if money supply shocks are highly persistent, their impact is mostly "anticipated" money growth, to which firms can adjust ahead of time. Therefore, the effects of the shock gradually tend to neutrality as persistence approaches one (a random walk in the money growth rate). However, persistence must be very high for this difference to be appreciable. The consumption and inflation responses to money growth shocks with persistence 0, 0.8, and 0.9 are almost indistinguishable. Even with $\phi_z = 0.99$, the cumulative consumption response is still around two-thirds of its level in the benchmark calibration of $\phi_z = 0.8$. Only when monthly persistence rises to 0.999 does the consumption response become trivial.[49]

3.6 Trend inflation

Finally, in Figures 9-10, we study how price dynamics in our nested logit model respond to large changes in inflation. We simulate trend inflation rates from -10% to 80% annually, which is the range of inflation documented by Gagnon (2009) in Mexican data from 1994 to 2002.[50]

FIGURE 9 ABOUT HERE

The left panel of Figure 9 shows how the frequency of price adjustment varies with trend inflation. Adjustment hazards are substantially higher in the Mexican data than the US data used in our calibration. We eliminate this level shift in the graph by normalizing the hazard to one at the observed 1994 pre-crisis inflation rate, in order to focus on changes in the frequency of adjustment. Our benchmark model matches the observed change in the hazard well: it roughly triples as inflation rises from 0% to 80%. Timing errors are important for this result. In the EiP model, where timing is frictionless, the hazard more than quadruples as inflation rises to 80%, a much larger change than that in the data.

The middle panel of the figure focuses instead on the size of price changes. Our model is somewhat less successful on this score: the average absolute price change doubles from its minimum of 8pp, to 16pp, as inflation rises to 80% annually, while in the Mexican data the size of price changes rises by 50%, to 12pp. On this statistic the EiP specification outperforms our benchmark model: the average size of price changes varies less with inflation if adjustments are affected by pricing errors but not by timing errors.

In the third panel, Figure 9 shows how the fraction of price increases varies with trend inflation. Our benchmark model closely tracks the fraction of increases in Mexican data, rising from 50% at 0% inflation to 93% at an 80% annual inflation rate. In contrast, the EiP, EiT, and FMC models all tend quickly to a limit with almost 100% price increases. Intuitively, the fat tails of the price histogram induced by the coexistence of pricing and timing errors help maintain the likelihood of price decreases, so 7% of price changes are negative even at the maximum inflation rate observed in the Mexican dataset.

FIGURE 10 ABOUT HERE

Since our model matches several features of price dynamics as trend inflation rises, it offers an interesting laboratory to analyze how trend inflation alters the effects of monetary shocks. Figure 10 compares the effects of a 1pp increase in money supply growth (with monthly auto-correlation 0.8) across different underlying trend inflation rates. The plain solid line represents our benchmark exercise, simulating the response of the nested model starting from 2% trend annual inflation; the figure compares how the effects differ at annual trend inflation rates of 0%, 10%, 20%, 40%, and 80%, respectively. The impact of money supply shocks on consumption goes rapidly towards zero, and becomes much less persistent, as the trend inflation rate rises above 10%. Qualitatively, this makes sense: the average adjustment frequency rises with inflation, making prices effectively less sticky, so nonneutrality decreases. Quantitatively, our model indicates that this effect is quite strong.

4 CONCLUSIONS

This paper has modeled nominal price rigidity as a near-rational phenomenon: price adjustment is costly because it requires the firm to spend time making decisions. We operationalize this idea by adopting the game-theoretic concept of "control costs" that increase with the precision of choice. We extend Mattsson and Weibull's (2002) result that decisions take the form of logit random variables when control costs depend on relative entropy to derive a logit hazard that governs the *timing* of price adjustment.

Our model implies that firms make errors on two margins– timing and pricing– and we study each margin separately to see how these two types of errors interact. The model with pricing errors but frictionless timing (EiP) implies that prices are sticky when they are near the optimum, since choice is costly and carries a risk of mistakes. These errors help match various features of retail price behavior, but the EiP setup lacks sufficient degrees of freedom to fit both the size distribution and the frequency of adjustments simultaneously. The model with errors in timing but frictionless price choices (EiT) generates substantially greater nonneutrality than EiP, because timing errors weaken the selection effect in adjustment dynamics. But by allowing for errors on both margins, the nested specification achieves a better fit to microdata while also reinforcing the nonneutrality of the EiT model.

Although we study a different source of nominal rigidity, our findings share some common ground with other recent studies of state-dependent pricing. Like Midrigan (2011), Álvarez *et al*. (2011), and Dotsey *et al*. (2013), we find that closely fitting a state-dependent pricing model to microdata implies greater monetary nonneutrality than a fixed menu cost model does, but less than that of a Calvo model with the same adjustment frequency. A similar mechanism is at work in all these papers: selection is weaker than in an FMC model because the hazard varies less abruptly with the value of adjustment — either because the firm is managing several prices simultaneously (Midrigan, 2011), or because it may not know the value of adjusting (Álvarez *et al*., 2011), or due to a prohibitively high draw of the stochastic menu cost (Dotsey *et al*., 2013). In the present paper, and in Costain and Nakov (2011B), we advocate an especially simple explanation for weak selection: small, low-cost errors of timing diminish the correlation between the value and the probability of adjustment.

While standard practice in microeconometrics includes error terms in all behavioral equations, most recent work on state-dependent pricing has instead modeled the full distribution of price changes as if firms' behavior were entirely error-free. Here, instead, we allow for mistakes, and interpret them structurally as the result of costly managerial decisions. The payoff is that by allowing for errors, we can eliminate all other frictions (including "menu costs" and exogenous probabilistic barriers to adjustment) but nonetheless match micro and macrodata at least as well as competing frameworks (many of which are less sparsely parameterized). While this paper has focused on price adjustment, our framework also seems applicable to a variety of contexts where a decision maker intermittently flips a switch or updates a number or a vector. Potential applications include wage bargaining, hiring and firing decisions, inventory control, portfolio adjustments, lumpy investment, and control of policy instruments.

References

[1] Álvarez, Fernando; Francesco Lippi; and Luigi Paciello. (2011) "Optimal price setting with observation and menu costs." *Quarterly Journal of Economics*, 126, 1909-1960.

[2] Álvarez, Fernando; Martín González-Rozada; Andy Neumeyer; and Martín Beraja. (2013) "From hyperinflation to stable prices: Argentina's evidence on menu cost models." SCID Working Paper No. 470, Stanford University.

[3] Anderson, Simon; André de Palma; and Jacques-François Thisse. (1992) *Discrete Choice Theory of Product Differentiation.* Cambridge, MA: MIT Press.

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

[5] Argente, David and Chen Yeh. (2017) "Product life cycle, learning, and nominal shocks." Manuscript, University of Chicago.

[6] Bachmann, Rudiger, and Giuseppe Moscarini. (2011) "Business cycles and endogenous uncertainty." 2011 Meeting Papers, Society for Economic Dynamics.

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

[8] Blanco, Andres. (2017) "Optimal inflation target in an economy with menu costs and zero lower bound." Manuscript, University of Michigan.

[9] Blundell, Richard, and Stephen Bond. (2000) "GMM Estimation with persistent panel data: an application to production functions." *Econometric Reviews*, 19, 321-340.

[10] Bono, James, and David Wolpert. (2009) "Statistical prediction of the outcome of a noncooperative game." American University Economics Working Paper No. 2009-20.

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

[12] Campbell, Jeffrey, and Benjamin Eden. (2014) "Rigid prices: evidence from U.S. scanner data." *International Economic Review*, 55, 423-442.

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

[14] Christiano, Lawrence; Martin Eichenbaum, and Charles Evans. (1999) "Monetary policy shocks: what have we learned and to what end?" In *Handbook of Macroeconomics*, Volume 1A, edited by John Taylor and Michael Woodford, pp. 65-148. Amsterdam: Elsevier.

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

[16] Costain, James, and Anton Nakov. (2011B) "Distributional dynamics under smoothly state-dependent pricing." *Journal of Monetary Economics*, 58, 646-65.

[17] Costain, James, and Anton Nakov. (2015A) "Precautionary price stickiness." *Journal of Economic Dynamics and Control*, 58, 218-34.

[18] Costain, James, and Anton Nakov. (2015B) "Logit price dynamics." CEPR Discussion Paper No. 10731.

[19] Costain, James. (2017) "Costly decisions and sequential bargaining." Banco de España Working Paper No. 1729.

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

[21] Demery, David. (2012) "State-dependent pricing and the non-neutrality of money." *Journal of Macroeconomics*, 34, 933-44.

[22] Dotsey, Michael; Robert King; and Alexander Wolman. (2013) "Inflation and real activity with firm-level productivity shocks." FRB Philadelphia Working Paper No. 13-35.

[23] Eichenbaum, Martin; Nir Jaimovich; and Sergio Rebelo. (2011) "Reference prices, costs and nominal rigidities." *American Economic Review*, 101, 234-262.

[24] Eichenbaum, Martin; Nir Jaimovich; Sergio Rebelo, and Josephine Smith. (2014) "How frequent are small price changes?" *American Economic Journal: Macroeconomics*, 6, 137-155.

[25] Gagnon, Etienne. (2009) "Price setting under low and high inflation: evidence from Mexico." *Quarterly Journal of Economics*, 124, 1221-63.

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

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

[28] Ilut, Cosmin; Rosen Valchev; and Nicolas Vincent. (2016) "Paralyzed by fear: Rigid and discrete pricing under demand uncertainty." NBER Working Paper No. 22490.

[29] Kehoe, Patrick, and Virgiliu Midrigan. (2015) "Prices are sticky after all." *Journal of Monetary Economics*, 75, 35-53.

[30] Khaw, Mel Win; Luminita Stevens; and Michael Woodford. (2017) "Discrete adjustment to a changing environment: Experimental evidence." *Journal of Monetary Economics*, 91, 88-103.

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

[32] Klenow, Peter, and Benjamin Malin. (2010) "Microeconomic evidence on price setting." In *Handbook of Monetary Economics*, Vol. 3, No. 3, edited by Benjamin Friedman and Michael Woodford, pp. 231-284. Amsterdam: Elsevier.

[33] Klein, Paul. (2000) "Using the generalized Schur form to solve a multivariate linear rational expectations model." *Journal of Economic Dynamics and Control*, 24, 1405-23.

[34] Machina, Mark. (1985) "Stochastic choice functions generated from deterministic preferences over lotteries." *Economic Journal*, 95, 575-94.

[35] Mackowiak, Bartosz, and Mirko Wiederholt. (2009) "Optimal sticky prices under rational inattention." *American Economic Review*, 99, 769-803.

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

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

[38] Matejka, Filip, and Alisdair McKay. (2015) "Rational inattention to discrete choices: a new foundation for the multinomial logit model." *American Economic Review*, 105, 272-298.

[39] Matejka, Filip. (2016) "Rationally inattentive seller: sales and discrete pricing." *Review of Economic Studies*, 83, 1156-1188.

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

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

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

[43] Midrigan, Virgiliu. (2011) "Menu Costs, Multi-Product Firms and Aggregate Fluctuations." *Econometrica*, 79, 1139-80.

[44] Nakamura, Emi, and Jón Steinsson. (2008) "Five facts about prices: a reevaluation of menu cost models." *Quarterly Journal of Economics*, 123, 1415-64.

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

[46] Shannon, Claude. (1948) "A mathematical theory of communication." *The Bell System Technical Journal*, 27, 379-423 and 623-56.

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

[48] Stahl, Dale. (1990) "Entropy control costs and entropic equilibrium." *International Journal of Game Theory*, 19, 129-38.

[49] Steiner, Jakub; Colin Stewart; and Filip Matejka. (2017) "Rational inattention dynamics: Inertia and delay in decision-making." *Econometrica*, 85, 521-553.

[50] Stevens, Luminita. (2015) "Coarse pricing policies." Federal Reserve Bank of Minneapolis Staff Report No. 520.

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

[52] Woodford, Michael. (2008) "Inattention as a source of randomized discrete adjustment." Manuscript, Columbia University.

[53] Woodford, Michael. (2009) "Information-constrained state-dependent pricing." *Journal of Monetary Economics*, 56, S100-S124.

[54] Yun, Tack. (2005) “Optimal Monetary Policy with Relative Price Distortions.” *American Economic Review*, 95, 89-109.

[55] 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, 514-33.

Notes

[1]For helpful comments, we thank H. Le Bihan, J. Campbell, A. Cheremukhin, R. Gary-Bobo, P. Karadi, A. Levin, F. Lippi, B. Mackowiak, F. Matejka, I. Méjean, G. Nuño, R. Reis, L. Stevens, A. Tutino, C. Walsh, M. Wiederholt, M. Woodford, and seminar participants at UC Santa Cruz, the Bank of Spain, CREST, FRB Richmond, the Bank of England, the ECB, BI Business School, Simposio AEE 2011, T2M 2012, ESSIM 2012, CEF 2012, EEA-ESEM 2012, the 2013 Chicago NBER Monetary Economics meeting, the 2013 BGSE Summer Forum, SED 2014, the 2014 NYU conference of the International Network on Expectations and Coordination, and the 2014 Vienna Macro Workshop, as well as the editor and referees. We also thank Borja Petit Zarzalejos (CEMFI) for excellent research assistance. We are grateful to Virgiliu Midrigan for making his data available, and to the Kilts Center at the Univ. of Chicago GSB, which is the source of those data. Views expressed here are those of the authors and may not coincide with those of the Bank of Spain, the Eurosystem, the ECB, or the CEPR.

[2]Modelling error-prone human behavior as a probability distribution over feasible actions has a long history; see Luce (1959), Machina (1985), or Anderson *et al.* (1992), Ch. 2.

[3]Logit equilibrium is a commonly-applied parametric special case of quantal response equilibrium (see McKelvey and Palfrey, 1995, 1998).

[4]This reflects much older results in physics, where a related optimization problem gives rise to the Boltzmann distribution of particles in a gas.

[5]We build on some of our previous work: Costain and Nakov (2015A) derived logit errors in prices, and one of the specifications in Costain and Nakov (2011A, B) imposed logit errors on adjustment timing.

[6]Several recent papers study a related source of retail price stickiness: learning about uncertain demand. See Bachmann and Moscarini (2011), Ilut, Valchev, and Vincent (2016), and Argente and Yeh (2017).

[7]The logit choice function is a well-known econometric framework for discrete choice. But logit *equilibrium*, in which each player makes logit decisions, based on payoff values consistent with other players' logit decisions, has to our knowledge rarely been applied outside of game theory.

[8]Entropy plays a role in Sims' approach, as a measure of information. It also enters our model, as an analytically convenient measure of precision.

[9]Much of the rational inattention literature, including Khaw *et al.* (2017) and Steiner *et al.* (2017), addresses partial equilibrium decision problems only. Alternatively, some rational inattention papers have computed general equilibrium by assuming a linear-quadratic-Gaussian economy or placing other strong restrictions on functional forms (Mackowiak and Wiederholt, 2009).

[10]These are gross real profits, prior to payment of the decision costs which we describe shortly.

[11]For concise notation, time subscripts on aggregate variables indicate dependence on the aggregate state. Thus, if Ξ_t represents the real aggregate state of the economy, we index equilibrium quantities by t to indicate that they are functions of Ξ_t. Hence $u_t(p, a) \equiv u(p, a, \Xi_t)$, and $C_t \equiv C(\Xi_t)$, and so forth.

[12]We could instead apply our model to intermittent updating of a real price, provided we are specific about the operational meaning of this control variable (*e.g.* indexing to some particular inflation indicator). To model control of a real price, we would alter the right-hand side of Bellman equation (14); for the empirical analysis, we would need data in which the real price changes could be observed or inferred.

[13]Allowing for a small "menu cost" too would be a simple but uninteresting extension; the model's behavior would be qualitatively unchanged.

[14] While our framework resembles Sims (2003), we assume precise choice is costly, given one's information, rather than assuming information itself is costly. We also depart from Sims (2003) by being specific about the control variable of the decision-maker, regarding each change in this control as a (costly) decision.

[15]Impatient readers who wish to skip the details of our decision cost function may jump directly to the two-stage representation of the solution, in Sec. 2.2.2.

[16]Since we will consider distributions that contain point masses, (10) is written as a Stieltjes integral instead of a Riemann integral. Moreover, at any x where $\pi_1(x)$ or $\pi_2(x)$ is nondifferentiable, the expression $\frac{\pi_1'(x)}{\pi_2'(x)}$ should be interpreted as $\lim_{\Delta \downarrow 0} \frac{\pi_1(x)-\pi_1(x-\Delta)}{\pi_2(x)-\pi_2(x-\Delta)}$. Specifically, if π_1 or π_2 contains a mass point at x, then $\frac{\pi_1'(x)}{\pi_2'(x)}$ represents the ratio of the point masses at x.

[17]More specific interpretations are possible. One might assume that the cognitive costs of controlling the price are paid primarily by a manager. Alternatively, even if managers can make error-free decisions costlessly, the firm will face control costs if employees must exert effort to implement management's instructions without error. In

general our cost structure could stand in for errors or miscommunications anywhere in the firm's management process. Absent conclusive empirical evidence we will not restrict our interpretation here.

[18]See our working paper, Costain and Nakov (2015B), for a nominal description of the model, which is then detrended it to obtain the real model studied here.

[19]In other words, for any nominal price P considered at time t, we have $\pi^\dagger(\ln(P/P_t)) = \Pi^\dagger(P)$ and $\theta^\dagger(\ln(P/P_t)) = \Theta^\dagger(P)$.

[20] Hence the set of nominal prices considered at time t is $\Gamma_t^P \equiv \{P : \ln(P/P_t) \in \Gamma^p\}$. But we do not actually need to construct Γ_t^P since we solve the real, detrended model instead of computing it in nominal terms.

[21]See Online Appendix C for an exploration of quantitative robustness under different cost functions.

[22]If instead the firm's control variable were its real price, then $i_{t,t+1}$ would not appear in (14).

[23]Cover and Thomas (2006), Theorem 2.7.2, shows that relative entropy is a nonnegative, convex function.

[24]Moreover, the mapping from the value function o_{t+1} to o_t is a contraction; see Costain (2017), Prop. 5. Therefore, the dynamic programming problem (13)-(14) converges to a unique solution.

[25] If instead $\theta^\dagger$ has a mass point at p, (15) determines the ratio of the mass points in $\pi^\dagger$ and $\theta^\dagger$, instead of the ratio of their densities.

[26]A hazard function of this form is also derived by Woodford (2008, 2009) from a model with a Shannon (1948) constraint.

[27]Matejka and McKay (2015) show that a static optimization problem with an entropy cost function, of form (20), is solved by the weighted logit (21).

[28]We thank Mike Woodford for suggesting this interpretation of the difference between our model and a rational inattention setup.

[29]Since economists are familiar with models of costless rationality, they often equate observing a given information set with knowing all quantities that can be calculated from it. But when reasoning is costly, these two assumptions are not equivalent. Here, we assume firms have enough information to calculate $v_t(p, a)$ if they wished. But since perfectly precise calculations are excessively costly in our framework, we do not equate possessing this information with actually knowing $v_t(p, a)$ or the optimal action.

[30]The operators d_a and $d_{\tilde{p}}$ indicate that integration is performed with respect to a and $\tilde{p}$, respectively.

[31]See Online Appendix C.2 for analysis of an economy governed by a Taylor rule. Our conclusions about state-dependent pricing are similar for the case of a Taylor rule to those we document here for the simple, transparent case of a money growth rule.

[32]The EiP case is discussed in greater detail in a previous paper, Costain and Nakov (2015A).

[33]See Figure 3, which compares these histograms in the data and in each specification of our model.

[34]Since the Euclidean norm of a vector scales with the square root of the number of elements, we scale the first term by $\sqrt{\#h}$ to place comparable weights on the two components of the distance measure.

[35]There are small quantitative changes if we double the support of θ, fixing other parameters. If we instead estimate the width of the support, jointly with κ and $\bar{\lambda}$, to minimize (33), then our results are visually indistinguishable from those reported here. These and other robustness results are detailed in Online Appendix C.

[36]The Calvo (1983) specification has a constant 10.2% adjustment hazard per month; the FMC setup requires firms to spend 0.0081 units of labor time to make a price change. These frameworks are otherwise frictionless, and are otherwise identical to our benchmark model.

[37]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 in our estimation criterion may not suffice to identify these parameters separately.

[38]The policy functions of the EiT and EiP limiting cases are illustrated in Figure 3 of our working paper, Costain and Nakov (2015B).

[39]Note that since we assume no menu costs— all costs relate to choice rather than price adjustment *per se*— firms have no reason to avoid a tiny price change if they calculate that this is the optimal thing to do.

[40]Eichenbaum *et al*. (2014) argue that many observations of small price changes are simply measurement error. However, little changes when we reestimate our model to take their finding into account. For example, if we alter our distance criterion (33) by assuming that all adjustments smaller than 1% are measurement error, our parameter estimate changes to $(\kappa, \bar{\lambda}) = (0.027, 0.21)$. The fit is slightly worse, since even our benchmark parameterization overpredicts small changes, but the micro and macro results we report are virtually unchanged.

[41]The table shows the expected gain in monthly revenues that would accrue to one infinitesimal firm if it could permanently make error-free decisions at zero cost, holding fixed the behavior of all other firms.

[42]Since the loss from nominal price stickiness is low, there is also little to be gained by setting a real price instead. The 1.41% revenue loss associated with setting a nominal price in the benchmark version falls to 1.36% if the firm's control variable is instead its real price.

[43]Since consumers are price takers in our model, all managerial costs of price adjustment in the model are related to decision-making rather than negotiation.

[44]The "reference" prices and costs reported by Eichenbaum *et al*. (2011) eliminate "sales" and other transitory changes. For their alternative measure of "weekly" prices and costs, the ratio of standard deviations is 1.08.

[45]For comparable graphs that show the EiP and EiT specifications too, see Costain and Nakov (2015B).

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

[47]Because the adjustment process is asymmetric (bottom panels of Figure 2), the steady state average error is nonzero. But time variation in the average error $\bar{\epsilon}_t$ is negligible.

[48]Hence the lines with dots in Figure 8 (showing the baseline value of persistence, $\phi_z = 0.8$), represent the integrals of the lines with dots in Figure 6.

[49]IRF calculations over longer time horizons confirm that the price level converges gradually to a 5 percentage point change, but that cumulative consumption plateaus within roughly ten months.

[50]While Gagnon (2009) studies relatively short-lived changes in the inflation rate, we compare his data to changes in trend inflation, which we can compute in a fully nonlinear way.

Figure 1: Timing of firm behavior.

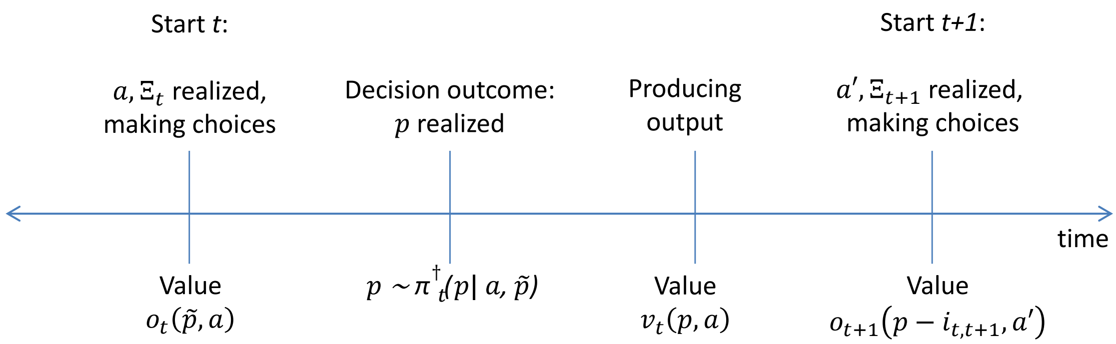
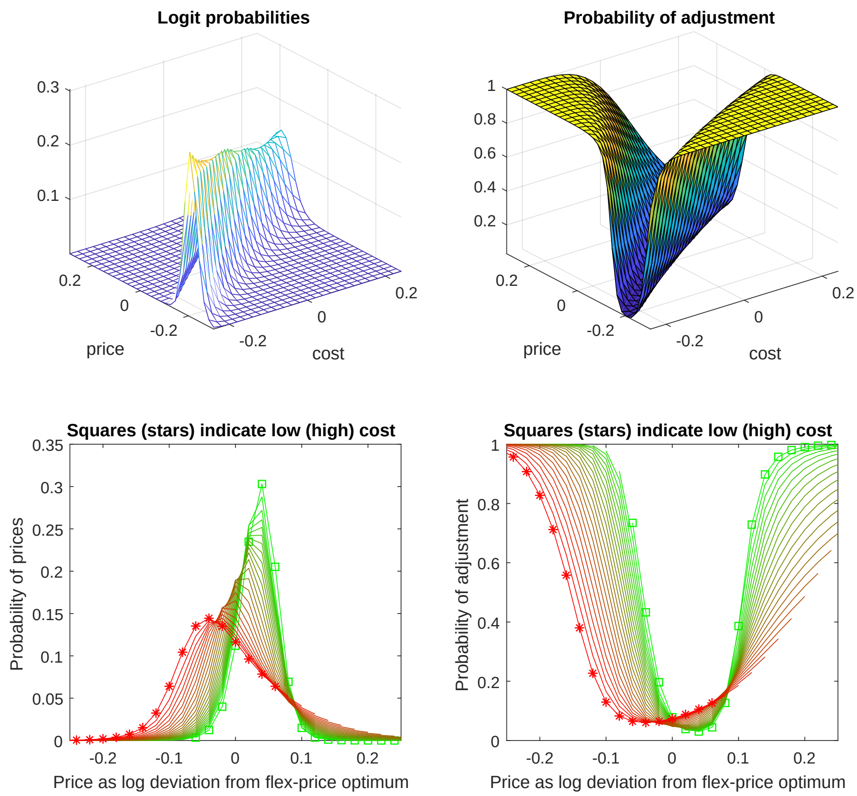


Figure 2: Price change distributions and adjustment function: nested model.



Notes:

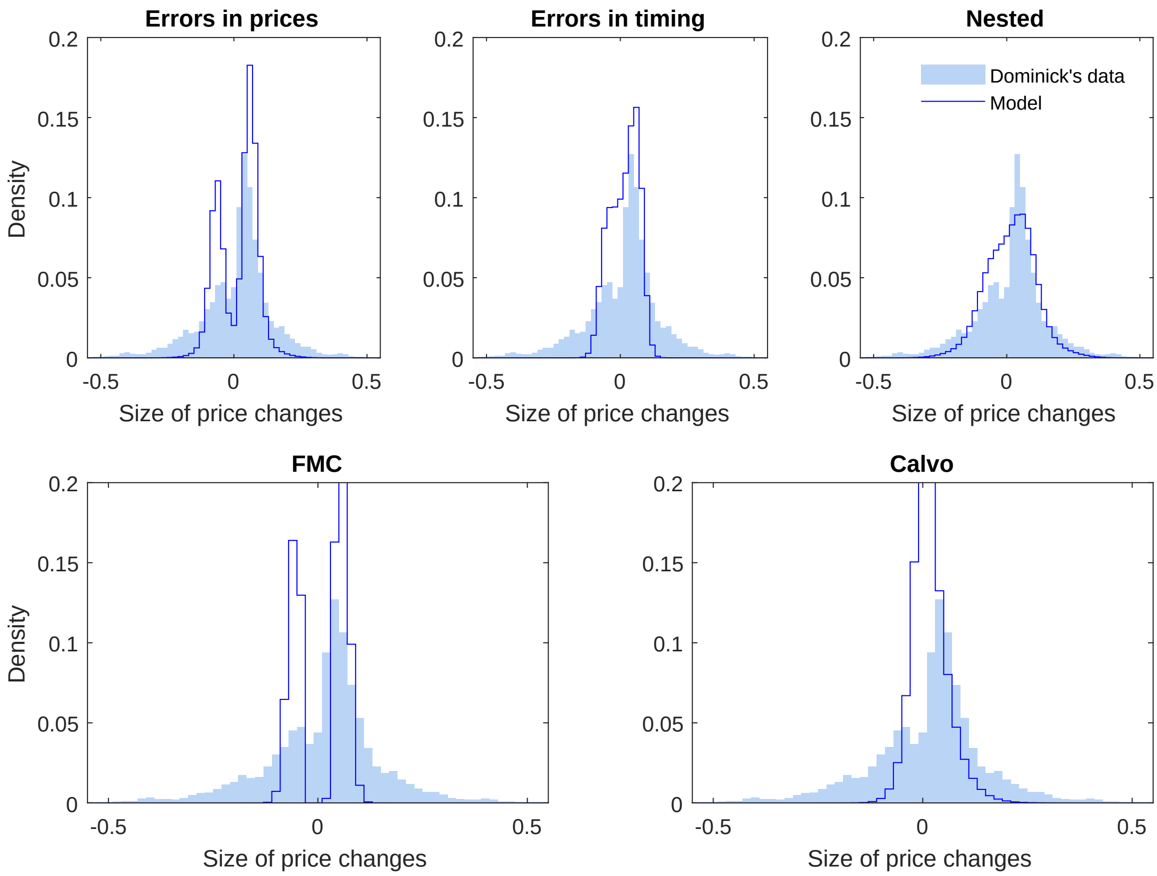
Prices and costs (A^{-1}) plotted as log deviations from their means.

Left panels: price distribution π conditional on cost.

Right panels: adjustment probability λ conditional on price and cost.

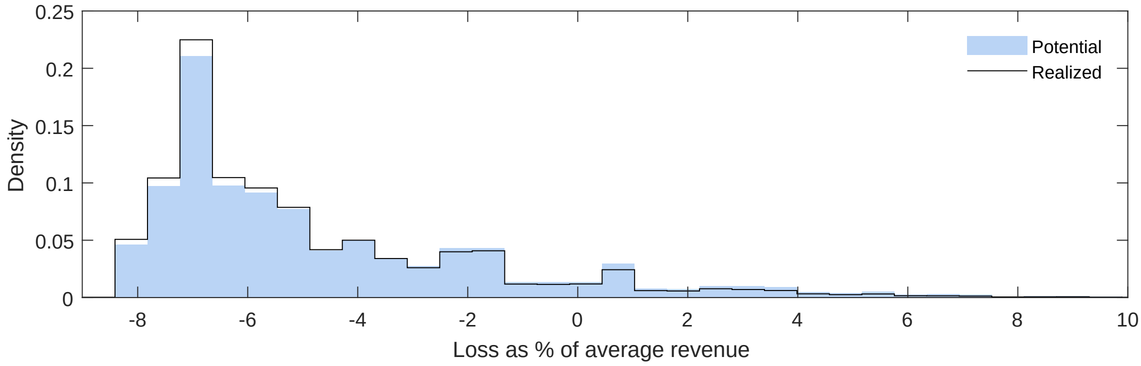
Lower panels plot slices at each cost A^{-1}, for $\ln A \in \Gamma^a$. Lowest cost highlighted with squares; highest cost highlighted with stars.

Figure 3: Distribution of price adjustments: comparing models.



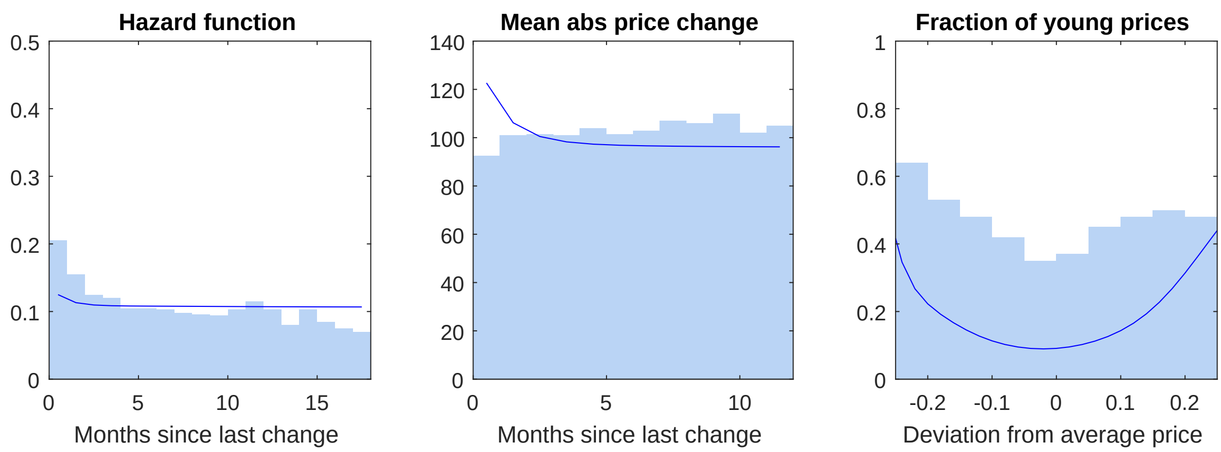
Notes:
Comparing histograms of nonzero log price adjustments.
Shaded area: Dominick's data. Lines: Model simulations.

Figure 4: Losses from failure to adjust: nested model.



Notes:
Histogram of losses $d(p,a) \equiv \tilde{v}(a) - v(p,a)$ from not adjusting, as a percentage of monthly average revenues. Potential losses before adjustments occur (with distribution $\widetilde{\Psi}(\tilde{p},a)$, shown as shaded area) and realized losses after adjustments ($\Psi(p,a)$, shown as solid line).

Figure 5: Price adjustment dynamics: nested model.



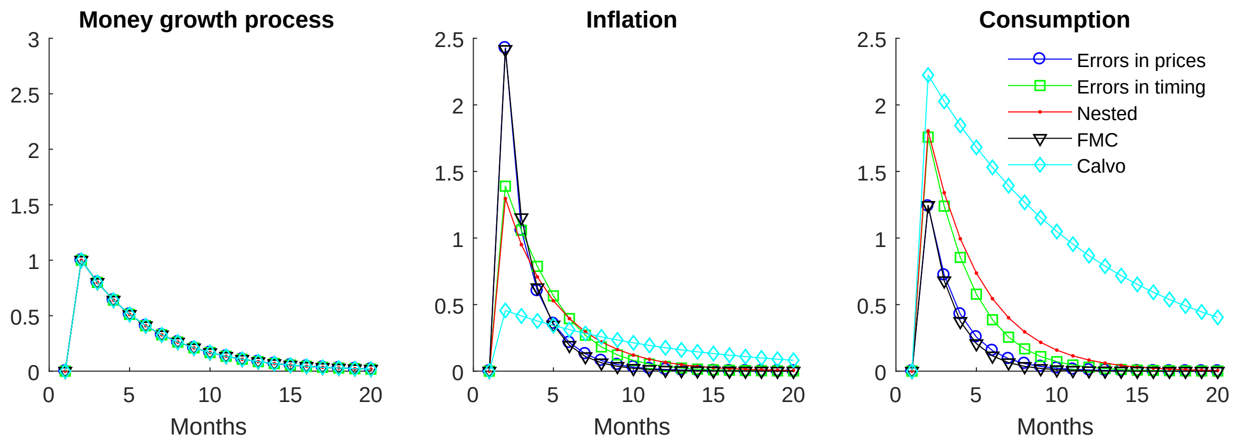
Notes:

Left: Monthly adjustment probability as function of time since last change. Data (shaded): Nakamura-Steinsson (2008).

Middle: Mean absolute change as function of time since last change. (Normalization: 100 = average adjustment.) Data (shaded): Klenow-Kryvtsov (2008).

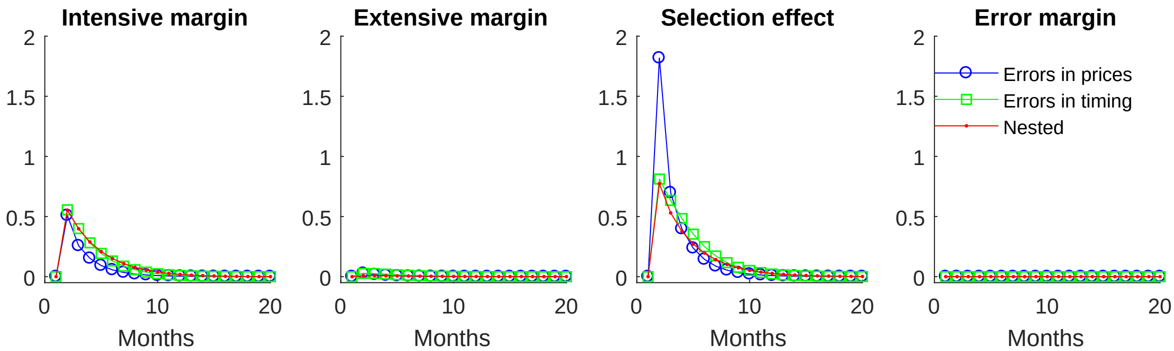
Right: Fraction of prices set within last two months, as function of deviation from average price in product class. Data (shaded): Campbell-Eden (2014).

Figure 6: Impulse responses to money growth shock: comparing models.



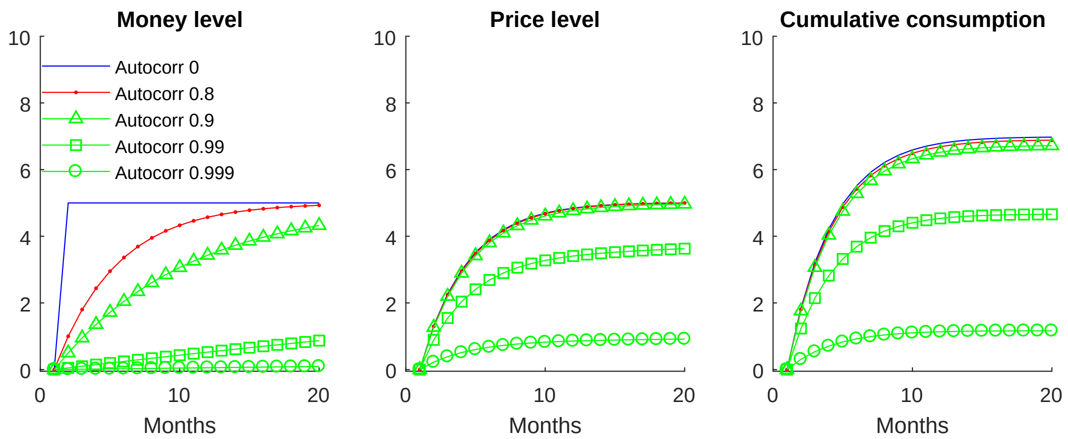
Notes:
Impulse responses of inflation and consumption to money growth shock with autocorrelation 0.8 (monthly). Circles: EiP. Squares: EiT. Dots: Nested. Triangles: FMC. Diamonds: Calvo.

Figure 7: Decomposition of inflation impulse responses: comparing models.



Notes:
Decomposition of inflation impulse response to money growth shock with autocorrelation 0.8 (monthly).
Circles: EiP version. Squares: EiT version. Dots: Nested version.

Figure 8: Impulse responses to money growth, as function of autocorrelation.

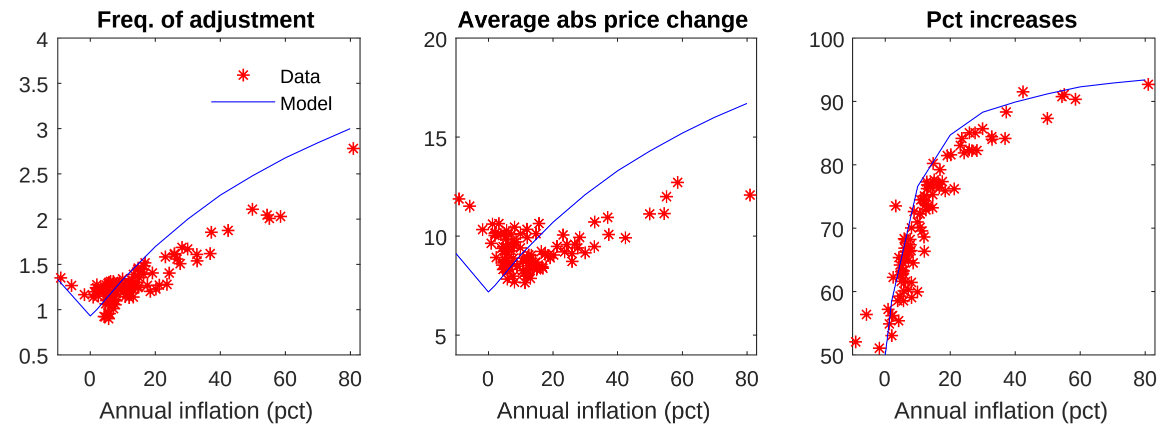


Notes:

Cumulative impulse responses of money stock, price level, and cumulative consumption in the nested model. Simple line: uncorrelated 5% money supply shock. Line with dots: 1% money shock with autocorrelation 0.8 (benchmark).

Triangles, squares, circles: autocorrelations 0.9, 0.99, and 0.999 monthly.

Figure 9: Effects of trend inflation: nested model.



Notes:

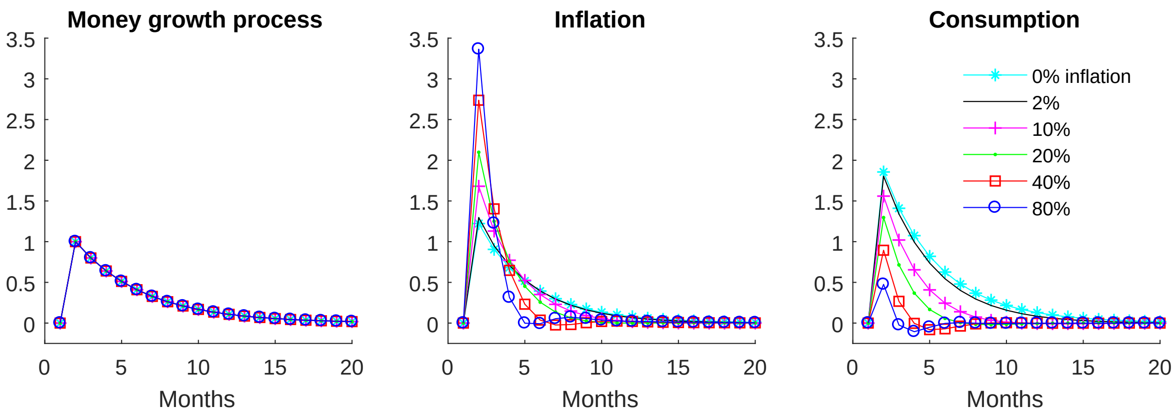
Left: Adjustment frequency as a function of trend inflation (normalized to one at 1994 mean inflation).

Middle: Average absolute price change (percent) as a function of trend inflation rate.

Right: Price increases as a percentage of price adjustments, as a function of trend inflation rate.

Stars: Gagnon (2009) Mexican data (monthly, March 1994 - May 2002). Lines: nested model.

Figure 10: Impulse responses to money growth shock, as function of trend inflation.



Notes:
Impulse responses of inflation and consumption to money shock with monthly autocorrelation 0.8, benchmark model. 1pp money growth shock, conditional on annual inflation rates of 0%, 2%, 10%, 20%, 40%, 80%.

Table 1: Adjustment parameters.*

	Calvo	FMC	Errors in timing (EiT)	Errors in prices (EiP)	Nested
$\bar{\lambda}$	–	–	0.045	–	0.22
κ_π	–	–	–	0.0044	0.018
κ_λ	–	–	0.0080	–	0.018
Exogenous hazard	0.102	–	–	–	–
Menu cost	–	0.0081	–	–	–

*Parameters are chosen to minimize (33); the "nested" model imposes $\kappa_\pi = \kappa_\lambda$.

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

	Calvo	FMC	Errors in timing	Errors in prices	Nested	Data
Adjustment frequency						
Frequency of price changes	10.2	10.2	10.2	10.2	10.2	10.2
Price change statistics						
Mean absolute price change	3.4	5.8	4.68	6.72	7.51	9.90
Std of price changes	4.4	5.8	5.27	7.32	9.30	13.2
Kurtosis of price changes	4.5	1.6	2.22	2.37	3.40	4.81
Percent of price increases	67.1	63.4	63.3	62.3	58.8	65.1
Percent of changes $\leq 5\%$	70.9	29.5	49.7	27.9	33.6	35.4
Variability of prices and costs						
$100 \times$ Std(p)/Std(a)	77.5	95.4	91.0	97.7	104	115**
Costs of decisions and errors						
Pricing costs*	0	0.19	0	0.174	0.509	
Timing costs*	0	0	0.167	0	0.361	
Loss relative to full rationality*	0.91	0.32	0.416	0.365	1.41	

Notes: 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, except for double asterisk, which indicates Eichenbaum *et al.* (2011).

Table 3: Variance decomposition and Phillips curves

Correlated money growth shock ($\phi_z = 0.8$)	Errors in timing	Errors in prices	Nested	Data
Freq. of price changes (%)	10.2	10.2	10.2	10.2
Std of money shock (%)	0.15	0.12	0.17	
Std of qtrly inflation (%)	0.25	0.25	0.25	0.25
% explained by μ shock alone	100	100	100	
Std of qtrly output growth (%)	0.37	0.20	0.43	0.51
% explained by μ shock alone	73	38	84	
Slope coeff. of Phillips curve	0.29	0.15	0.35	
R^2 of regression	0.94	0.85	0.98	

Notes: 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 money supply growth rate and superscript q indicates quarterly averages.
Dataset: Dominick's.

Online Appendices for “Logit price dynamics”, July 2018

James Costain (Banco de España and European Central Bank)

Anton Nakov (European Central Bank and CEPR)

ONLINE APPENDIX A: COMPUTATION

Outline of algorithm

Heterogeneity is a challenge when computing our model: at any time t, productivities A_{it} and prices P_{it} will differ across firms. The Calvo model is popular because, up to a first-order approximation, only the average price matters for equilibrium. But this property does not hold in most sticky-price models, in which equilibrium quantities depend on the whole time-varying distribution of prices and productivity across firms.

To address this issue, we apply Reiter's (2009) solution method for dynamic general equilibrium models with heterogeneous agents and aggregate shocks. As a first step, the algorithm calculates the steady-state general equilibrium in the absence of aggregate shocks. Idiosyncratic shocks are still active, but are assumed to have converged to their ergodic distribution, so the real aggregate state of the economy is a constant, Ξ. The algorithm solves a discretized approximation of the underlying model; here we restrict real log prices p_{it} and log productivities a_{it} to 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 both uniformly spaced (in logs). We can then view the steady state value function as a matrix $\mathbf{V}$ of size $\#^p \times \#^a$, comprising the values $v^{jk} \equiv v(p^j, a^k, \Xi)$ associated with prices and productivities $\left(p^j, a^k\right) \in \Gamma$.[1] 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 end of any given period. To calculate steady state general equilibrium, we can guess the wage w, then

[1] In this appendix, bold face indicates matrices, and superscripts represent indices of matrices or grids.

solve the firm's problem by backwards induction on the grid Γ, then update the conjectured wage, and iterate to convergence.

The second step 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 value function is represented by a $\#^p \times \#^a$ matrix $\mathbf{V}_t$ with row j, column k element $v_t^{jk} \equiv v(p^j, a^k, \Xi_t)$, summarizing the time t values at all grid points $(p^j, a^k) \in \Gamma$. Then, instead of treating the Bellman equation as a functional equation that defines $v(p, a, \Xi)$ for all possible idiosyncratic and aggregate states p, a, and Ξ, we view it as a difference equation linking the matrices $\mathbf{V}_t$ and $\mathbf{V}_{t+1}$. This amounts to a (large!) system of $\#^p\#^a$ first-order expectational difference equations governing the $\#^p\#^a$ variables v_t^{jk}. We linearize these equations numerically (together with the $\#^p\#^a$ equations that govern the distribution $\mathbf{\Psi}_t$, and a few other scalar equations). We solve the linearized model using the QZ decomposition, following Klein (2000).

This method combines linearity and nonlinearity in a way appropriate for models of price setting, where idiosyncratic shocks tend to be more relevant for firms' decisions than aggregate shocks are. By linearizing the aggregate dynamics, we recognize that changes in the aggregate shock z_t or in the distribution $\mathbf{\Psi}_t$ are unlikely to have a highly nonlinear impact on the value function. This smoothness does not require any "approximate aggregation" property, in contrast with the Krusell and Smith (1998) method; nor do we need to impose any particular functional form on the distribution $\mathbf{\Psi}$. However, to allow for the strong impact of firm-specific shocks, the method treats variation along idiosyncratic dimensions in a fully nonlinear way: the value at each grid point is determined by a distinct equation.

The discretized model

In the discretized model, the value $\mathbf{V}_t$ is a $\#^p \times \#^a$ matrix with elements $v_t^{jk} \equiv v(p^j, a^k, \Xi_t)$ for $\left(p^j, a^k\right) \in \Gamma$. A uniform default distribution θ allocates probability $1/\#^p$ to each price in Γ^p. Solving a single Bellman step analytically, the expected value of setting a new price is a row vector $\tilde{\mathbf{v}}_t$ of length $\#^a$, with kth element

$$\tilde{v}_t^k \;\equiv\; \kappa_\pi w_t \ln\left(\frac{1}{\#^p}\sum_{j=1}^{\#^p}\exp\left(\frac{v_t^{jk}}{\kappa_\pi w_t}\right)\right). \tag{40}$$

The value function $\mathbf{O}_t$ is also a $\#^p \times \#^a$ matrix, as is the hazard policy $\mathbf{\Lambda}_t$ and the logit price probabilities policy $\mathbf{\Pi}_t$; their (j,k) elements are given by[2]

$$o_t^{jk} \equiv \kappa_\lambda w_t \left(\bar{\lambda}\exp\left(\frac{\tilde{v}_t^k}{\kappa_\lambda w_t}\right) + (1-\bar{\lambda})\exp\left(\frac{v_t^{jk}}{\kappa_\lambda w_t}\right)\right), \tag{41}$$

$$\lambda_t^{jk} \equiv \bar{\lambda}\left(\bar{\lambda} + (1-\bar{\lambda})\exp((v_t^{jk} - \tilde{v}_t^k)/(\kappa w_t))\right)^{-1}, \tag{42}$$

$$\pi_t^{jk} \equiv \frac{\exp\left(v_t^{jk}/(\kappa w_t)\right)}{\sum_{n=1}^{\#p}\exp\left(v_t^{nk}/(\kappa w_t)\right)}. \tag{43}$$

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

In this discrete representation, the productivity process (34) can be written as a $\#^a \times \#^a$ matrix $\mathbf{S}$, where the (m,k) element represents the following transition probability:

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

Likewise, we can write the impact of inflation on real prices in Markovian notation. Let $\mathbf{R}_t$ be a $\#^p \times \#^p$ matrix in which element (m,l) represents the probability that firm i's beginning-of-t

[2]Equation (42) is a simplified description of λ_t^{jk}. While (42) implies that λ_t^{jk} represents the function $\lambda_t(p^j, a^k)$ evaluated at the log price grid point p^j and log productivity grid point a^k, in our computations λ_t^{jk} in fact represents the *average* of $\lambda_t(\tilde{p}, a^k)$ 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 $v_t(\tilde{p}, a^k)$ between price grid points. Defining λ_t^{jk} this way ensures differentiability with respect to changes in the aggregate state Ω_t.

log real price $\tilde{p}_{it}$ equals $p^m \in \Gamma^p$, if its log real price at the end of $t-1$ was $p^l \in \Gamma^p$:

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

Generically, the deflated log price $p_{i,t-1} - i_{t-1,t}$ will fall between two grid points; then the matrix $\mathbf{R}_t$ must round up or down stochastically.[3] Also, if $p_{i,t-1} - i_{t-1,t}$ lies below the smallest or above the largest element of the grid, then $\mathbf{R}_t$ must round up or down to keep prices on the grid.[4] Unbiased rounding results if $\mathbf{R}_t$ is constructed as:

$$R_t^{ml} = prob(\widetilde{p}_{it} = p^m | p_{i,t-1} = p^l, i_t) = \begin{cases} 1 & \text{if } p^l - i_t \leq p^1 = p^m \\ \frac{p^l - i_t - p^{m-1}}{p^m - p^{m-1}} & \text{if } p^1 < p^m = \min\{p \in \Gamma^p : p \geq p^l - i_t\} \\ \frac{p^{m+1} - p^l + i_t}{p^{m+1} - p^m} & \text{if } p^1 \leq p^m = \max\{p \in \Gamma^p : p < p^l - i_t\} \\ 1 & \text{if } p^l - i_t > p^{\#^p} = p^m \\ 0 & \text{otherwise.} \end{cases} \tag{44}$$

The distributional dynamics can now be written in compact matrix form; eq. (23) becomes:

$$\widetilde{\mathbf{\Psi}}_t = \mathbf{R}_t * \mathbf{\Psi}_{t-1} * \mathbf{S}', \tag{45}$$

where $*$ represents ordinary matrix multiplication. Productivity shocks are represented by right multiplication, while transitions in the real price level are represented by left multiplication. Next, to calculate the effects of price adjustment on the distribution, let $\mathbf{E}_{pp}$ and $\mathbf{E}_{pa}$ be matrices of ones of size $\#^p \times \#^p$ and $\#^p \times \#^a$, respectively. Eq. (24) is then:

$$\mathbf{\Psi}_t = (\mathbf{E}_{pa} - \mathbf{\Lambda}) \,.* \,\widetilde{\mathbf{\Psi}}_t + \mathbf{\Pi}_t \,.* \,(\mathbf{E}_{pp} * (\mathbf{\Lambda} \,.* \,\widetilde{\mathbf{\Psi}}_t)), \tag{46}$$

[3] If instead the firm's control variable were its real price, then $\mathbf{R}_t$ would simply be an identity matrix.

[4] In other words, any nominal price leading to a real log price below p^1 after inflation is automatically rounded up to the real log price p^1 (and to compute examples with deflation we must shift down any real log price exceeding $p^{\#^p}$). This assumption is made for numerical purposes only, and has a negligible impact on the equilibrium as long as Γ^p is sufficiently wide.

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

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

$$u_t^{jk} \equiv \left(\exp(p^j) - \frac{w_t}{\exp(a^k)}\right)\frac{C_t}{\exp(\epsilon p^j)} \tag{47}$$

for $\left(p^j, a^k\right) \in \Gamma$. Then Bellman equation (14) becomes:

$$\mathbf{V}_t = \mathbf{U}_t + \beta E_t\left\{\frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}}\left[\mathbf{R}'_{t+1} * \mathbf{O}_{t+1} * \mathbf{S}\right]\right\}. \tag{48}$$

The expectation E_t in (48) refers only to the effects of the time $t+1$ aggregate shock z_{t+1}, because the expectation over idiosyncratic states $(p^j, a^k) \in \Gamma$ is represented by multiplying by $\mathbf{R}'_{t+1}$ and $\mathbf{S}$. Note that since (48) iterates backwards in time, its transitions are governed by $\mathbf{R}'$ and $\mathbf{S}$, whereas (45) iterates forward in time, involving $\mathbf{R}$ and $\mathbf{S}'$.

We now discuss how we apply Reiter's (2009) two-step method to this discrete model.

Step 1: steady state

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

Long run monetary neutrality implies that nominal money growth rate equals the inflation rate in steady state: $\mu = \exp(i)$. Thus, the steady-state transition matrix $\mathbf{R}$ is known, since it depends only on inflation i, and the Euler equation reduces to $\exp(i) = \beta R$.

We can then calculate general equilibrium as a one-dimensional root-finding problem in w.

Given w, we calculate $C = (w/\chi)^{1/\gamma}$, and then construct matrix $\mathbf{U}$, with elements

$$u^{jk} \equiv \left(\exp(p^j) - \frac{w}{\exp(a^k)}\right) \frac{C}{\exp(\epsilon p^j)} \,. \tag{49}$$

We can then find the fixed point of the value $\mathbf{V}$ (simultaneously with $\tilde{\mathbf{v}}$ and $\mathbf{O}$):

$$\mathbf{V} = \mathbf{U} + \beta \mathbf{R}' * \mathbf{O} * \mathbf{S}. \tag{50}$$

This allows us to calculate the logit matrix $\mathbf{\Pi}$, with elements

$$\pi^{jk} \equiv \frac{\exp\left(v^{jk}/(\kappa w)\right)}{\sum_{n=1}^{\#p} \exp\left(v^{nk}/(\kappa w)\right)}. \tag{51}$$

Likewise, we calculate the hazard matrix $\mathbf{\Lambda}$. We can then find the steady state distribution by iterating on the two-step distributional dynamics:

$$\mathbf{\Psi} = (\mathbf{E}_{pa} - \mathbf{\Lambda}) \,.*\, \widetilde{\mathbf{\Psi}} + \mathbf{\Pi} \,.*\, (\mathbf{E}_{pp} * (\mathbf{\Lambda} \,.*\, \widetilde{\mathbf{\Psi}})) \tag{52}$$

$$\widetilde{\mathbf{\Psi}} = \mathbf{R} * \mathbf{\Psi} * \mathbf{S}' \tag{53}$$

Finally, we check whether

$$1 \;=\; \sum_{j=1}^{\#^p} \sum_{k=1}^{\#^a} \Psi^{jk} \exp\left((1-\epsilon)p^j\right) \equiv p(w) \tag{54}$$

If $p(w) = 1$, then an equilibrium value of w has been found.

Step 2: linearized dynamics

Given the steady state, the general equilibrium dynamics can be calculated by linearization. First, we eliminate as many variables from the equation system as we can, summarizing the dynamics in terms of the exogenous shock process z_t, the lagged distribution of idiosyncratic states $\mathbf{\Psi}_{t-1}$, and the endogenous "jump" variables including $\mathbf{V}_t$, $\mathbf{\Pi}_t$, C_t, m_{t-1}, and i_t. The equation system reduces to

$$z_t = \phi_z z_{t-1} + \epsilon_t^z \tag{55}$$

$$\frac{\mu \exp(z_t)}{\exp i_t} = \frac{m_t}{m_{t-1}} \tag{56}$$

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

$$\mathbf{V}_t = \mathbf{U}_t + \beta E_t \left\{ \frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} \left[\mathbf{R}'_{t+1} * \mathbf{O}_{t+1} * \mathbf{S} \right] \right\} \tag{58}$$

$$1 = \sum_{j=1}^{\#^p} \sum_{k=1}^{\#^a} \Psi_t^{jk} \exp((1-\epsilon)p^j) \tag{59}$$

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

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

then the whole set of expectational difference equations (55)-(59) governing the dynamic equilibrium becomes a first-order system of the following form:

$$E_t \mathcal{F}\left(\overrightarrow{X}_{t+1}, \overrightarrow{X}_t, z_{t+1}, z_t \right) = 0 \tag{60}$$

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

To see that the vector $\overrightarrow{X}_t$ in fact contains all the variables we need, note that given i_t and i_{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 $\mathbf{\Psi}_{t-1}$. Given $w_t = \chi C_t^{\gamma}$, we can construct $\mathbf{U}_t$, with (j,k) element equal to $u_t^{jk} \equiv \left(\exp(p^j) - \frac{w_t}{\exp(a^k)} \right) \frac{C_t}{\exp(\epsilon p^j)}$. Finally, given $\mathbf{V}_t$, and $\mathbf{V}_{t+1}$ we can construct $\mathbf{\Pi}_t$ and $\tilde{\mathbf{v}}_t$, and thus $\mathbf{\Lambda}_t$ and $\mathbf{O}_{t+1}$. Therefore the variables in $\overrightarrow{X}_t$ and z_t are indeed sufficient to evaluate the system (55)-(59).

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 a linear first-order expectational difference equation system:

$$E_t\mathcal{A}\Delta\overrightarrow{X}_{t+1} + \mathcal{B}\Delta\overrightarrow{X}_t + E_t\mathcal{C}z_{t+1} + \mathcal{D}z_t = 0 \tag{61}$$

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.

ONLINE APPENDIX B: SEQUENTIAL STATEMENT OF THE OPTIMIZATION PROBLEM

In this appendix, we discuss a sequential representation of the firm's partial-equilibrium decision problem. In this representation, we think of the firm as choosing a plan contingent on any possible history up to a given time T. In particular, it must consider histories $(\tilde{p}^T, a^T, \Xi^T)$ incorporating its own prices and productivity, and aggregate states, where superscripts indicate time series: $\tilde{p}^T \equiv (\tilde{p}_0, \tilde{p}_1, \ldots, \tilde{p}_T)$, and likewise for a^T and Ξ^T.

By repeated substitution in the recursive problems (13)-(14), we can derive the following sequential optimization problem:

$$o_0(\tilde{p}_0, a_0) \quad = \quad \max_{\pi_t^\dagger \in \Delta(\Gamma^p(\tilde{p}_t))} E_0 \sum_{t=0}^{\infty} q_{0,t} \left[\int u_t(p_t, a_t) d\pi_t^\dagger(p_t) - \kappa w_t \mathcal{D}\left(\pi_t^\dagger || \theta^\dagger(\cdot|\tilde{p}_t)\right) \right] \tag{62}$$

$$\text{s.t. } \tilde{p}_t = p_{t-1} - i_{t-1,t}, \quad \text{and} \quad \int d\pi_t^\dagger(p) = 1 \ \text{ for all } t. \tag{63}$$

Here E_0 refers to an expectation calculated under the dynamics of the firm's productivity shock a_t and the dynamics of the aggregate state Ξ_t. The discount factor is $q_{0,t} \equiv \Pi_{s=0}^{t-1} q_{s,s+1}$, where $q_{t,t+1} \equiv \beta \frac{P_t C_{t+1}^{-\gamma}}{P_{t+1} C_t^{-\gamma}}$; we assume discount factors satisfy $\sum_{t=0}^{\infty} q_{0,t} < \infty$.

The choice problem here is to be understood as choosing a function $\pi_t^\dagger(\tilde{p}^t, a^t, \Xi^t)$ conditional on each history $(\tilde{p}^t, a^t, \Xi^t)$ of length t. Here $\tilde{p}_t = p_{t-1} - i_{t-1,t}$ represents the log real price at the beginning of t, prior to the choice of $\pi_t^\dagger$, and $\Delta(\Gamma^p(\tilde{p}_t))$ is the set of increasing functions f satisfying $f(\min \Gamma^p(\tilde{p}_t)) \geq 0$; constraint (63) ensures that f is a *c.d.f.* Notice that if two functions π_t^1 and π_t^2 both lie in the set $\Delta(\Gamma^p(\tilde{p}_t))$ and both integrate to one over $\Gamma^p(\tilde{p}_t)$, then so does any convex combination of those functions. The same argument can be made for history-contingent plans. Therefore we conclude that the choice set in problem (62)-(63) is convex.

Moreover, the objective function contains two terms (for each t): expected profits, which are a linear function of $\pi_t^\dagger$, minus decision costs, which are a convex function of $\pi_t^\dagger$. Therefore the objective function of (62)-(63) is concave.

We should also be careful to check that the constraint set of problem (62)-(63) is non-empty, and that the objective function is finite-valued. Note that the strategy $\pi_t^\dagger = \theta^\dagger(\cdot|\tilde{p}_t)$ is feasible, and has zero decision cost, attaining the value

$$\underline{o}_0(\tilde{p}_0, a_0) \;\equiv\; E_0 \sum_{t=0}^{\infty} \int q_{0,t} u_t(p_t, a_t) d\theta^\dagger(p_t|\tilde{p}_t) \;>\; -\infty.$$

This lower bound is finite because sets Γ^a and Γ^p are assumed bounded, and the profit function in (8) is continuous. On the other hand, the value of (62)-(63) is bounded above by

$$\overline{o}_0(a_0) \;\equiv\; E_0 \sum_{t=0}^{\infty} q_{0,t} \max_p u_t(p, a_t) \;<\; \infty.$$

Therefore the value of (62) is bounded: $\underline{o}_0(\tilde{p}_0, a_0) < o_0(\tilde{p}_0, a_0) < \overline{o}_0(a_0)$.

Thus (62)-(63) maximizes a concave function over a non-empty, convex set, attaining a finite value. Hence there can be at most one solution to the first-order conditions, and if such a solution is found, it represents a solution to the optimization problem (62)-(63). Indeed, the first-order conditions yield the same solution that we found for the recursive problem (13)-(14).[5]

[5] Deriving the first-order conditions of (62)-(63) is tedious, so we omit them here, but they closely resemble those of (13)-(14). Where the value $v_t(p_t, a_t)$ appears in (15), we instead find a term of the form $E_t \sum_{s=0}^{\infty} q_{t,t+s} prob(p_{t+s} = p_t - i_{t,t+s}) u_{t+s}(p_t - i_{t,t+s}, a_{t+s})$, where $i_{t,t+s} \equiv \ln(P_{t+s}/P_t)$. This term represents the discounted sum of profits at all future times $t+s$ conditional on the nominal price set at time t remaining unadjusted at time $t+s$.

ONLINE APPENDIX C: ROBUSTNESS OF THE RESULTS

C.1 Changing the decision cost function

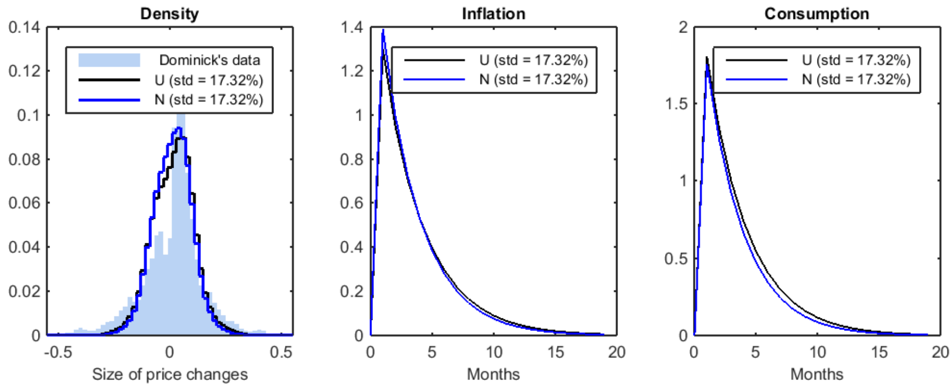
In our model, the precision of price choices is measured by comparing the firm's chosen price distribution to an exogenously fixed "default" distribution. We have run extensive simulations to explore whether our results are robust to changes in the assumed default distribution. In summarizing our conclusions, it is useful to distinguish two properties of the default distribution – its functional form, and its standard deviation. We find that our results are qualitatively and quantitatively robust to changes in both of these properties. There is a simple reason for this: we estimate that decision costs are low. Since deviating from the default distribution is not very costly, the precise form of that default has little impact on our results.

If we generalize the uniform default probabilities assumed in our benchmark parameterization, then the weighted logit (21) no longer reduces to the unweighted logit (43). To see how the results differ, compare the blue and black lines in the price change histogram (left panel) and impulse responses (middle and right panels) shown in Figure C.1. The black lines represent the benchmark uniform specification; the blue lines instead assume a truncated normal default distribution, with the same standard deviation. The results are almost identical. We have computed several other examples which show that the form of the default distribution has very little effect. In other words, firms' optimization, represented by $\exp(v/(\kappa w))$ in the logit formula (21), is powerful enough that reweighting by a different distributional form θ hardly matters.

The fact that we define the default distribution on a discrete grid is likewise irrelevant for the results. Computation on a discrete grid is a matter of numerical necessity. However, making this grid much finer has entirely negligible effects, both on the steady state and on the dynamic implications of the model.

Another change that might seem especially relevant would be to allow the default distribution to vary over time by recentering it on the previous nominal price. We simulated a speci-

Figure C.1: Uniform versus normal default distributions.



Notes:
Comparing benchmark uniform default distribution (as in paper) with truncated normal default distribution.
Left panel: Histogram of nonzero log price adjustments (Dominick's data shown as blue bars).
Middle and right panels: impulse responses to money growth shock with monthly autocorrelation 0.8.

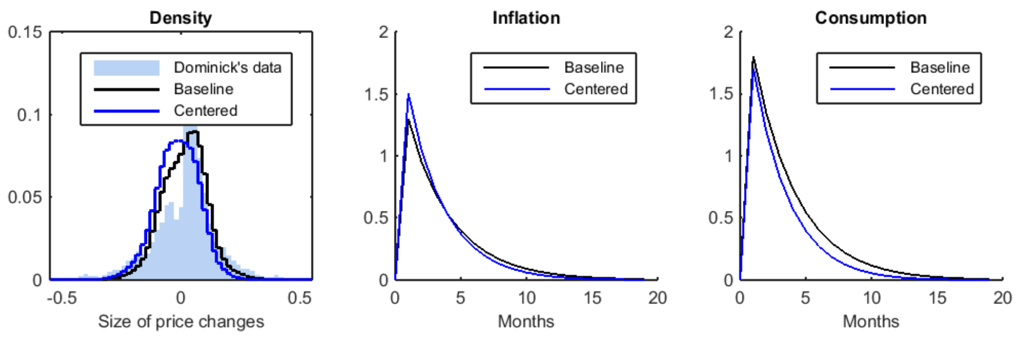
fication of this type, assuming a truncated normal default $\theta(p|\tilde{p})$ centered around $\tilde{p}$, the price prior to adjustment. Figure C.2 compares this specification to the unchanging uniform distribution used in our benchmark calculations.[6] The recentered normal specification (blue) implies a somewhat more symmetric histogram, a small increase in the adjustment hazard, and a resulting small decrease in the real effects of a money shock. But overall the differences are minor.

The effects of increasing the standard deviation of the default distribution are shown in Figure C.3. The benchmark results are shown in black; the effects of making the grid 100% wider (while fixing the grid step size) are shown in green.[7] Doubling the grid width makes the histogram more strongly bimodal (improving fit in the center while making it worse in the tails); the frequency of adjustment decreases from 10.2% to 6.6% monthly (because choosing from a wider range of prices amounts to a more difficult decision problem) and therefore the real effects of the money shock increase slightly.

Thus, changing the standard deviation of the default distribution has a small but nontrivial

[6]Figure C.2, like Figure C.1, changes the form of the default distribution without altering its standard deviation.
[7]We have also studied the effects of increasing the standard deviation of the default distribution when the default is a truncated normal. The (small) effects are similar to those shown in Figure C.3 for the uniform case.

Figure C.2: Uniform versus recentered normal default distributions.

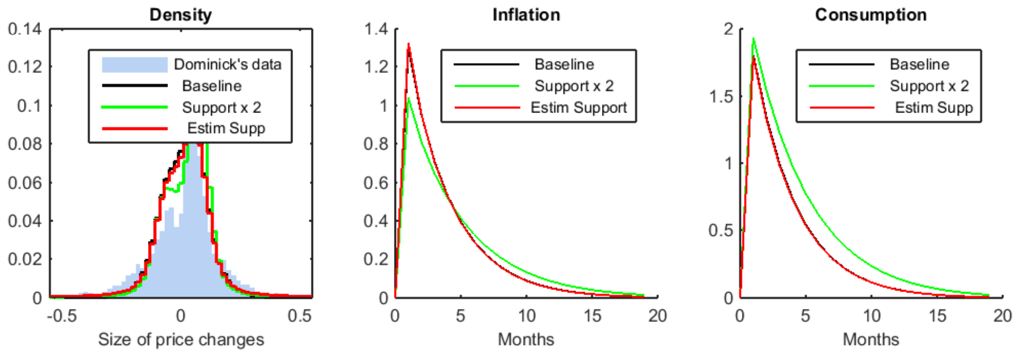


Notes:
Comparing benchmark uniform default distribution (as in paper) with truncated normal default distribution $\theta(p|\tilde{p})$ centred around pre-adjustment price $\tilde{p}$.
Left panel: Histogram of nonzero log price adjustments (Dominick's data shown as blue bars).
Middle and right panels: impulse responses to money growth shock with monthly autocorrelation 0.8.

impact on the results. But would it make any difference if we treated the standard deviation of the default as another free parameter to be estimated? This question is also addressed in Figure C.3, where the red line shows the results of jointly estimating κ, $\bar{\lambda}$, and the width of the price grid Γ^p to maximize our estimation criterion (33). The estimation favors a harder decision problem than we assumed in our benchmark calibration (the preferred grid is slightly more than twice as wide as the benchmark grid) but this is compensated by a slightly less error-prone and substantially quicker decision process ($\kappa = 0.17$ and $\bar{\lambda} = 0.35$, in contrast to the previous values $\kappa = 0.18$ and $\bar{\lambda} = 0.22$). The implied adjustment frequency is again 10.2% monthly, with the result that the impulse responses are almost indistinguishable from those in the benchmark specification (the black benchmark IRF is almost invisible under the red IRF resulting from estimating the width of the grid; likewise the black and red price adjustment histograms are almost identical). So while widening the grid, *ceteris paribus*, slightly increases monetary nonneutrality, estimating the grid width jointly with our other parameters gives results nearly identical to our benchmark parameterization.

Why are the impulse responses unchanged when we reestimate the model? All of the robust-

Figure C.3: Widening the support of the uniform default distribution.



Notes:
Black: uniform default distribution (benchmark from paper). Green: uniform default with 100% wider support.
Red: reestimating model with width of support of default as a free parameter. (Grid step size fixed.)
Left panel: Histogram of nonzero log price adjustments (Dominick's data shown as blue bars).
Middle and right panels: impulse responses to money growth shock with monthly autocorrelation 0.8.

ness exercises that we have run suggest that as long as our model matches the 10.2% adjustment hazard of our estimation criterion, the degree of monetary nonneutrality is virtually unchanged. Obviously this does not mean that *all* models with a 10.2% adjustment hazard are equivalent; the Calvo model implies much larger real effects, as our paper shows. But our error-prone model is very robust to changes in the specification of the default distribution, as long we parameterize the model to fit the estimation criterion.

To summarize, changing the shape of the default distribution is not quantitatively relevant for our results. Neither is treating its standard deviation as a free parameter to be estimated. What matters is that price decisions are somewhat noisy (helping fit the microdata) and timing decisions are also somewhat noisy (diminishing the selection effect and generating higher nonneutrality). The precise form of the noise is not at all essential for these conclusions.

C.2 Extending the model

Throughout the paper we have studied a stripped-down general equilibrium structure in order to focus primarily on the role of price stickiness. However, our framework can readily be extended to incorporate a more complete macroeconomic environment. Building a full medium-scale DSGE model is beyond the scope of this paper, but in this section we consider two especially relevant extensions. Our main conclusions about state-dependent nominal rigidity are unaltered.

On one hand, there is no need to restrict monetary policy to a money growth rule. Here we instead consider a Taylor-style interest rate rule of the form

$$i_t = \phi_i i_{t-1} + (1-\phi_i)\phi_\pi \left[\pi_t - \ln(\mu)\right] + \epsilon_t^i,$$

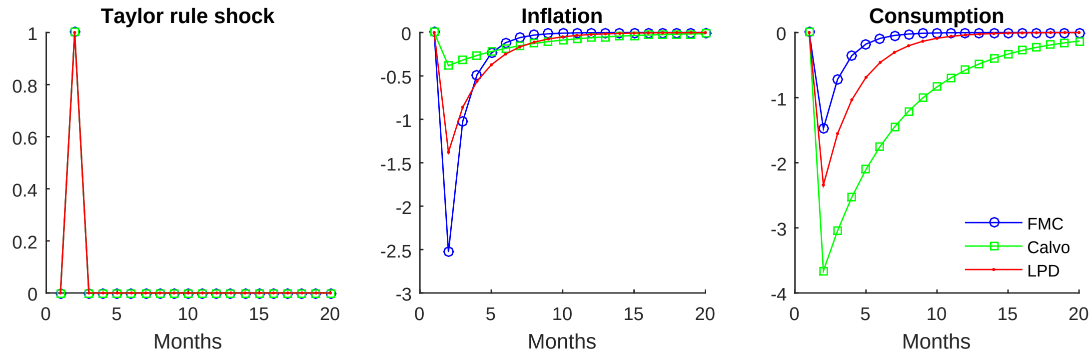
where i_t is the net nominal interest rate, π_t is the net inflation rate; μ is the steady-state target inflation rate, ϕ_i is an interest rate smoothing parameter; ϕ_π controls the strength of monetary policy reaction to inflation; and ϵ_t^i is an interest rate shock. This rule replaces the money supply equation (26).

Second, while nominal rigidity is central to generating real responses to purely nominal disturbances, it is likely that multiple forms of real rigidity also play a role in propagating shocks. In particular, Blanco (2017) imposes a production function with intermediate inputs, as a realistic and tractable source of real rigidities that can reinforce the effects of nominal rigidity. Here we modify the production function as follows:

$$Y_{it} = A_{it} N_{it}^{1-\eta} M_{it}^{\eta},$$

where M_{it} denotes the intermediate inputs used in the production of the differentiated final

Figure C.4: Impulse responses to a Taylor rule shock, with intermediate inputs in production.



Notes:
Impulse responses of inflation and consumption to an *i.i.d.* interest rate shock (percentage points). Comparing models of nominal rigidity.

goods Y_{it}. The goods market clearing condition becomes

$$Y_t = C_t + \int M_{it} di \, ;$$

C_t is now replaced by Y_t in equation (31).

We set $\phi_i = 0.9$, $\phi_\pi = 2$, and $\eta = 1/3$. The impulse responses to a monetary policy shock are shown in Figure C.4, which compares our nested benchmark specification (marked LPD), with FMC and Calvo models that likewise allow for a Taylor rule and real rigidities. The main purpose of this exercise is simply to show that the central results of our paper go through in this extended version. Namely, the nested specification produces real effects that fall between those of FMC and Calvo, and the relative degree of nonneutrality across these frameworks is quantitatively similar to what we found previously. When calibrating a model for applied purposes, real rigidities and a more realistic description of monetary policy are relevant elements to include in the analysis. But the difference in nonneutrality implied by our framework, relative to alternative models nominal rigidity, appears robust to these extensions.