ACCOUNTING FOR PERSISTENCE AND VOLATILITY OF GOOD-LEVEL REAL EXCHANGE RATES: THE ROLE OF STICKY INFORMATION

by

Mario J. Crucini, Mototsugu Shintani, and Takayuki Tsuruga

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

Accounting for persistence and volatility of good-level real exchange rates: the role of sticky information[*]

Mario J. Crucini[†] Mototsugu Shintani[‡] and Takayuki Tsuruga[§]

First Draft: October 2007

This version: April 2008

Abstract

Volatile and persistent real exchange rates are observed not only in aggregate series but also in the individual good level data. Kehoe and Midrigan (2007) recently showed that, under a standard assumption on nominal price stickiness, empirical frequencies of micro price adjustment cannot replicate the time-series properties of the law-of-one-price deviations. We extend their sticky price model by combining good specific price adjustment with information stickiness in the sense of Mankiw and Reis (2002). Under a reasonable assumption on the money growth process, we show that the model fully explains both persistence and volatility of the good-level real exchange rates. Furthermore, our framework allows for multiple cities within a country. Using a panel of U.S.-Canadian city pairs, we estimate a dynamic price adjustment process for each 165 individual goods. The empirical result suggests that the dispersion of average time of information update across goods is comparable to that of average time of price adjustment.

JEL Classification: E31, F31, D40

Keywords: Good-level real exchange rates, Law of one price, Sticky information, Dynamic panel

[†]Department of Economics, Vanderbilt University

[‡]Department of Economics, Vanderbilt University and Institute for Monetary and Economic Studies, Bank of Japan

[§]Faculty of Economics, Kansai University

1

1 Introduction

Aggregate real exchange rates are among the most scrutinized of economic variables because their persistence and volatility are much higher than what economists believe is consistent with a plausible degree of price rigidity. The time-dependent pricing model offers a convenient theoretical framework linking price stickiness and real exchange rate behavior. Chari, Kehoe, and McGrattan (2002, CKM) show that to generate the observed persistence of CPI-based aggregate real exchange rates, prices need to be exogenously fixed for at least one year. This degree of price-stickiness, however, appears implausible based on recent evidence of Bils and Klenow (2004) who find a median duration between price changes of only 4.3 months in U.S. micro-data.

An emerging literature using international micro-data finds the half-life of deviations from the Law of One Price (LOP) for the median good in the neighborhood of 18 months, considerably lower than the consensus 3-5 year half-lives of aggregate real exchange rates (Crucini and Shintani (2008)). This evidence suggests that studies using prices of individual goods, rather than price indices, is a more promising approach for evaluating time-dependent pricing models and under- standing short-run international relative price dynamics. An important contribution along this line is Kehoe and Midrigan (2007) who allow different price stickiness across individual goods and show that the persistence in LOP deviations is equal to 'the Calvo parameter,' the probability of price non-adjustment at the good level. Their empirical analysis using real exchange rates of 66 individual goods shows that the frequency of no price adjustment is higher for goods that also exhibit more persistence deviations from the LOP, as suggested by the theoretical model. How- ever, the persistence puzzle is still not resolved in the sense that the observed frequencies of micro price changes are too high to replicate the persistence of real exchange rate for most goods in the cross-section. In addition, the model does not match the time series variability of LOP deviations observed in the micro-data. These theoretical and empirical results are important and point to the need to break the tight correspondence between the frequency of price adjustment and the LOP persistence parameter characterizing the standard Calvo-type sticky price model.

In this paper, we break this tight link by extending the Kehoe-Midrigan model to allow for information stickiness. That is, in addition to the standard Calvo pricing, we assume only a fraction of firms update their information set each month, with the fraction possibly differing across firms. Thus price dynamics become a convolution of price adjustment timing and information updating. In the macroeconomic literature, Mankiw and Reis (2002) show that a model of information stickiness,

or inattentiveness, is capable of explaining the observed slow response of aggregate inflation to monetary shocks much better than sticky prices alone. When the information stickiness augments the Calvo-type sticky price mechanism, less frequent information updating leads to higher price persistence, at a given frequency of price adjustment (Dupor, Kitamura, and Tsuruga (2008, DKT)). With plausible assumptions on the money growth processes of two countries in the international setting, a similar effect takes place to increase both the persistence and volatility of real exchange rates.

In addition to the generalization of the sticky price model to allow for the information stickiness, our analysis differs from Kehoe and Midrigan (2007) in several aspects. First, our empirical analysis is based on an international retail price survey which records local currency prices for highly disaggregated individual goods and services spanning most of the CPI basket. Using this survey we expand the number of products from 66 products used in Kehoe and Midrigan (2007) to 165. Another advantage of this data is that the survey is conducted by a single agency, the Economist Intelligence Unit, so we can expect a reasonable uniformity in the quality of the products among international cities. An important limitation of our data is its annual frequency and relative short time-span, from 1990 to 2005. As in the case of Crucini and Shintani (2008), the difficulty of estimating persistence with short time-series is mitigated by utilizing the dynamic panel feature of the data.

Second, our theoretical model allows for the presence of multiple cities in each country and for long-run price deviations between the cross-border city pairs to differ by good and city pair. For each good, we use the panel of 52 U.S.-Canadian city pairs to estimate a dynamic panel model and to compute the volatility under the error components model framework.

Third, we also examine the effect of the exclusion of sales on the performance of sticky price models in explaining real exchange rate dynamics. Recently, Nakamura and Steinsson (2007) claim that the evidence of the fast price adjustment obtained by Bils and Klenow (2004) may be strongly influenced by the presence of sales, or other temporary price reductions. Nakamura and Steinsson (2007) define the regular price change by excluding sales from the observed price change, and report that the median frequency of regular price changes increases to the range of 8 to 11 months. Since prices are stickier based on this alternative definition of price change, it elevates the Calvo model's ability to account for important features of the data. We evaluate the performance of the model using both of these definitions of price adjustment frequency.

The main conclusions of Kehoe and Midrigan (2007) are robust to the change in the data. We

confirm that the both persistence and volatility are much higher than the prediction of a standard Calvo-type sticky price model even if we use (i) more disaggregated retail price data, (ii) panel data consisting of multiple cities in the U.S. and Canada, and (iii) adjust the frequency of price changes for temporary sales.

However, unlike the standard Calvo model, our extended model with information stickiness can fully account for both persistence and volatility. The model fares well when the average duration between information updates is 14 to 17 months if sales are not removed and 9 to 12 months if sales are removed. The ability of our model in fully replicating the observed persistence and volatility contrasts to another possible extension of the Calvo model allowing for pricing complementarities. Kehoe and Midrigan (2007) show that such an extension only leads to a modest improvement in explaining the persistence and little improvement in explaining the variance. Our key finding is that the dispersion of average duration between information updates across goods is comparable to the average duration between price changes. While the existing micro evidence on information stickiness is quite limited relative to that on price stickiness, our estimates of the information delay parameter seem consistent with the available survey evidence on the frequency with which firms conduct major information updates. Given the cost of information collection and processing for pricing decisions, we view information stickiness story as realistic as the traditional menu cost story.[1]

This paper is organized as follows: Section 2 presents our model as a generalization of Kehoe and Midrigan's model. In Section 3 we examine the model implications for the time series properties of the good-level real exchange rates. Section 4 describes our data and how we use it to evaluate the model. We also compare the benchmark sticky price model and our extended model. The study ends with a discussion of future research in Section 5 .

2   The model

Trade is over a continuum of goods between two countries with multiple cities located in each country. Under monopolistic competition, firms set prices in local currency to satisfy demand for a particular good in a particular city. A representative agent in each country chooses consumption over an infinite horizon subject to a cash-in-advance (CIA) constraint. In what follows, the U.S. and Canada represent the home and foreign country, respectively, and the unit of time is one month.

---

[1]Examples includes Sims (2003), Woodford (2003) and Mankiw and Reis (2006).

The lowest level of aggregation is the brand, z of a particular good. U.S. brands of each good are indexed $z \in [0, 1/2]$ while those in Canada are indexed $z \in (1/2, 1]$. Integrating over brands, we have the CES indices for consumption of good j in a U.S. city l and a Canadian city l^*, given by

$$c_t(j, l) = \left[\int_0^1 c_t(j, l, z)^{\frac{\theta-1}{\theta}} dz \right]^{\frac{\theta}{\theta-1}} \tag{1}$$

and

$$c_t^*(j, l^*) = \left[\int_0^1 c_t^*(j, l^*, z)^{\frac{\theta-1}{\theta}} dz \right]^{\frac{\theta}{\theta-1}}, \tag{2}$$

where $c_t(j, l, z)$ is consumption of a brand z of good j in U.S. city denoted l and $c_t^*(j, l^*, z)$ is the analog consumption of that brand for a Canadian city, l^*.

CES aggregation across cities $l \in [0, 1]$ and $l^* \in [0, 1]$, gives national consumption of good j within the U.S.

$$c_t(j) = \left[\int c_t(j, l)^{\frac{\theta-1}{\theta}} dl \right]^{\frac{\theta}{\theta-1}}, \tag{3}$$

and Canada,

$$c_t^*(j) = \left[\int c_t^*(j, l^*)^{\frac{\theta-1}{\theta}} dl^* \right]^{\frac{\theta}{\theta-1}}, \tag{4}$$

respectively.

CES aggregation across goods in each country gives aggregate consumption in the U.S., c_t,

$$c_t = \left[\int c_t(j)^{\frac{\theta-1}{\theta}} dj \right]^{\frac{\theta}{\theta-1}} \tag{5}$$

and Canada, c_t^*,

$$c_t^* = \left[\int c_t^*(j)^{\frac{\theta-1}{\theta}} dj \right]^{\frac{\theta}{\theta-1}}. \tag{6}$$

2.1 Households

As in Kehoe and Midrigan (2007) complete markets for state-contingent money claims exist. Agents decide how many one-period nominal bonds to hold in each state of the world in period $t+1$. U.S. households hold B_{t+1} while Canadians hold B_{t+1}^* (both denominated in the U.S. dollars).[2] The price of a bond issued at date t, maturing at date $t+1$ is denoted by $Q_{t,t+1}$. Also, $Q_{t,t+h}$ is the nominal stochastic discount factor by which all firms, regardless of their country of origin, discount profits earned in period $t+h$ back to the present period t.

---

[2] As Kehoe and Midrigan (2007) argue, it does not matter if foreign (Canadian) consumers hold complete and state-contingent one-period nominal bonds denominated in the foreign currency (Canadian dollars). It would be simply a redundant assumption under state-contingent bond markets.

Households in the each country maximize the discounted sum of $U(c_t, n_t) = \ln c_t - \chi n_t$ $(\chi > 0)$ subject to an intertemporal budget constraint and a CIA constraint. The maximization problem for U.S. households is:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(c_t, n_t), \tag{7}$$

$$\text{s.t.} \quad M_t + \mathbb{E}_t(Q_{t,t+1}B_{t+1}) = R_{t-1}W_{t-1}n_{t-1} + B_t + (M_{t-1} - P_{t-1}c_{t-1}) + T_t + \Pi_t, \tag{8}$$

$$M_t \geq P_t c_t, \tag{9}$$

where β is the discount factor of the household satisfying $0 < \beta < 1$ and $\mathbb{E}_t(\cdot)$ denotes the expectation operator conditional on the information available in period t.

The left hand side of the intertemporal budget constraint (8) represents the nominal value of total wealth of the household brought into the beginning of period $t+1$. It consists of cash holding M_t and bond holdings B_{t+1}. As shown in the right hand of (8), the household receives nominal labor income $W_{t-1}n_{t-1}$ in period $t-1$ which earns gross nominal interest R_{t-1} per unit of labor income until period t in terms of U.S. currency.[3] Households carry nominal bonds in amount B_t and cash holding remaining after consumption expenditures $(M_{t-1} - P_{t-1}c_{t-1})$ into period t; P_t is the aggregate price index defined below. Finally, T_t and Π_t are nominal lump sum transfers from the U.S. government and nominal profits of firms operating in the U.S., respectively.[4]

The equation (9) is the CIA constraint. The aggregate price P_t is given by $P_t = [\int P_t(j)^{1-\theta} dj]^{\frac{1}{1-\theta}}$, where $P_t(j)$ is the aggregate price index for good j; it is a CES aggregate over city-specific prices for that good: $P_t(j) = [\int P_t(j,l)^{1-\theta} dl]^{\frac{1}{1-\theta}}$. The price index for good j in a particular city l used in this aggregation is given by

$$P_t(j,l) = \left[\int P_t(j,l,z)^{1-\theta} dz \right]^{\frac{1}{1-\theta}}.$$

Households in Canada solve the analogous optimization problem except we must convert their U.S. dollar bond holdings into Canadian dollars at the spot nominal exchange rate, S_t. Thus the Canadian-dollar intertemporal budget constraint is

$$M_t^* + \frac{\mathbb{E}_t(Q_{t,t+1}B_{t+1}^*)}{S_t} = \frac{S_{t-1}R_{t-1}}{S_t}W_{t-1}^* n_{t-1}^* + \frac{B_t^*}{S_t} + (M_{t-1}^* - P_{t-1}^* c_{t-1}^*) + T_t^* + \Pi_t^*.$$

---

[3]We assume that the government pays interest rate $R_t(= 1/\mathbb{E}_t Q_{t,t+1})$ on labor income in period t. This assumption allows households' intratemporal optimality condition to be undistorted.

[4]We assume that government's lump sum transfers and firms' profits in a country go to households in that country.

The first order conditions of households in both countries are as follows:

$$\frac{W_t}{P_t} = \chi c_t \tag{10}$$

$$\frac{W_t^*}{P_t^*} = \chi c_t^* \tag{11}$$

$$\mathbb{E}_t Q_{t,t+1} = \beta \mathbb{E}_t \left[\left(\frac{c_{t+1}}{c_t} \right)^{-1} \frac{P_t}{P_{t+1}} \right] \tag{12}$$

$$\mathbb{E}_t Q_{t,t+1} = \beta \mathbb{E}_t \left[\left(\frac{c_{t+1}^*}{c_t^*} \right)^{-1} \frac{S_t P_t^*}{S_{t+1} P_{t+1}^*} \right] \tag{13}$$

$$M_t = P_t c_t \tag{14}$$

$$M_t^* = P_t^* c_t^*. \tag{15}$$

The equations (10) and (11) represent intratemporal substitution between labor and consumption while (12) and (13) represent intertemporal consumption choices across adjacent months. The intertemporal conditions, (12) and (13), are slightly different because Canadians buy state-contingent one-period nominal bonds denominated in the U.S. dollars. The CIA constraints always bind as shown in equations (14) and (15).

The nominal wage rate in a country is proportional to the stock of money held by households in that country. Combining the intratemporal conditions (10) and (11) with the CIA constraints we have:

$$W_t = \chi M_t, \tag{16}$$

$$W_t^* = \chi M_t^*. \tag{17}$$

The aggregate real exchange rate is determined by combining the home and foreign intertemporal conditions:

$$q_t = \frac{S_t P_t^*}{P_t} = \kappa \frac{c_t}{c_t^*}, \tag{18}$$

where $\kappa = q_0 c_0^* / c_0$.[5]

The nominal exchange rate is determined by combining (18) with the CIA constraints (14) and (15):

$$S_t = \kappa \frac{M_t}{M_t^*}. \tag{19}$$

---

[5]See Appendix A.

2.2 Firms

The output of brand z of good j in the U.S. is equal to the number of hours allocated to that activity:

$$y_t(j,z) = n_t(j,z) \ . \tag{20}$$

Goods are perishable, so the consumption of each good across all cities equals output of that good in the current period:

$$\int c_t(j,l,z)dl + \int (1+\tau(j,l^*))c_t^*(j,l^*,z)dl^* = y_t(j,z). \tag{21}$$

We allow for long-run deviations from the LOP across borders through $\tau(j,l^*)$, an iceberg transportation cost in exporting good j from the U.S. to a Canadian city indexed by l^*. A firm must ship $(1+\tau(j,l^*))$ units of good j to city l^* for one unit of that good to arrive at the destination.

An analogous market clearing condition holds for each of the Canadian goods:

$$\int (1+\tau(j,l))c_t(j,l,z)dl + \int c_t^*(j,l^*,z)dl^* = y_t^*(j,z). \tag{22}$$

2.3 Price adjustment and information updating

This section begins by reviewing Calvo pricing used by Kehoe and Midrigan (2007) and then presents our extension to allow for information updating as in Mankiw and Reis (2002). The equilibrium is briefly described in each setting.

2.3.1 Calvo pricing

We model the nominal price rigidities as in Calvo (1983) and Yun (1996): each month a fraction of firms $1 - \lambda_j$ are randomly drawn and allowed to reset their prices. As suggested by the subscript, the frequency of price changes varies according to the type of good j and is assumed to be the same in both countries, good-by-good.

All U.S. firms that sell their good j in city l choose the same optimal price when they adjust prices in period t. The price $P_{H,t}(j,l)$ solves the following maximization problem:

$$\max_{P_{H,t}(j,l)} \mathbb{E}_t \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h}[P_{H,t}(j,l) - W_{t+h}]$$
$$\times \left(\frac{P_{H,t}(j,l)}{P_{t+h}(j,l)}\right)^{-\theta} \left(\frac{P_{t+h}(j,l)}{P_{t+h}(j)}\right)^{-\theta} \left(\frac{P_{t+h}(j)}{P_{t+h}}\right)^{-\theta} c_{t+h}, \tag{23}$$

for all cities $l \in [0,1]$. Here, we used the three demand functions as constraints:

$$c_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\theta} c_t$$

$$c_t(j,l) = \left(\frac{P_t(j,l)}{P_t(j)}\right)^{-\theta} c_t(j)$$

$$c_t(j,l,z) = \left(\frac{P_t(j,l,z)}{P_t(j,l)}\right)^{-\theta} c_t(j,l).$$

The optimality condition for $P_{H,t}(j,l)$ is

$$
\mathbb{E}_t \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h} \left(\frac{P_{H,t}(j,l)}{P_{t+h}}\right)^{-\theta} c_{t+h}
$$
$$
= \frac{\theta}{\theta-1} \mathbb{E}_t \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h} \left(\frac{W_{t+h}}{P_{H,t}(j,l)}\right) \left(\frac{P_{H,t}(j,l)}{P_{t+h}}\right)^{-\theta} c_{t+h}. \tag{24}
$$

Similarly, all Canadian firms that export and sell their good j in city l choose the same optimal price $P_{F,t}(j,l)$ when they adjust prices. The price $P_{F,t}(j,l)$ for these firms solves the maximization problem:

$$
\max_{P_{F,t}(j,l)} \mathbb{E}_t \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h} [P_{F,t}(j,l) - (1 + \tau(j,l)) S_{t+h} W_{t+h}^*]
$$
$$
\times \left(\frac{P_{F,t}(j,l)}{P_{t+h}(j,l)}\right)^{-\theta} \left(\frac{P_{t+h}(j,l)}{P_{t+h}(j)}\right)^{-\theta} \left(\frac{P_{t+h}(j)}{P_{t+h}}\right)^{-\theta} c_{t+h}, \tag{25}
$$

for all cities $l \in [0,1]$. The optimality condition is of the form similar to (24):

$$
\mathbb{E}_t \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h} \left(\frac{P_{F,t}(j,l)}{P_{t+h}}\right)^{-\theta} c_{t+h}
$$
$$
= \frac{\theta}{\theta-1} \mathbb{E}_t \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h} \left(\frac{(1+\tau(j,l)) S_{t+h} W_{t+h}^*}{P_{F,t}(j,l)}\right) \left(\frac{P_{F,t}(j,l)}{P_{t+h}}\right)^{-\theta} c_{t+h}. \tag{26}
$$

2.3.2 Calvo pricing with infrequent information updating

We now add information stickiness following Mankiw and Reis (2002) to the model. Consider firms facing two nominal rigidities. First, each firm has a constant probability of price resetting $1 - \lambda_j$ as before. Second, with probability of $1 - \omega_j$, a firm receives an information update in the current month. The fraction of firms that fail to get updates, ω_j, use the information available from the most recent update. For tractability, we assume that the two probabilities are independent each other.

DKT develop this combined stickiness structure to explain persistent inflation dynamics as we specified above. In DKT, infrequent price changes arise due to the Calvo assumption of price

changes. However, when firms compute their optimal reset prices, a fraction of firms use the newest information set and the remaining firms use the stale information set to determine prices. Following DKT, we employ this structure and refer to it as "dual stickiness" pricing.

All U.S. firms that sell their good j in city l choose different prices according to the vintage of their information set. When firms are allowed to adjust prices, those with the same vintage of information choose the same price. Let $P_{H,t}^k(j,l)$ be the optimal reset price set by U.S. firms conditional on information of vintage k, its age in months. The price $P_{H,t}^k(j,l)$ for these firms solves

$$\max_{P_{H,t}^k(j,l)} \mathbb{E}_{t-k} \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h}[P_{H,t}^k(j,l) - W_{t+h}]$$
$$\times \left(\frac{P_{H,t}^k(j,l)}{P_{t+h}(j,l)} \right)^{-\theta} \left(\frac{P_{t+h}(j,l)}{P_{t+h}(j)} \right)^{-\theta} \left(\frac{P_{t+h}(j)}{P_{t+h}} \right)^{-\theta} c_{t+h},$$

(27)

for $k = 0, 1, 2, \cdots$ and for all cities $l \in [0,1]$. Note the only difference between this problem and the standard Calvo problem is that the expectation is taken with respect to information of vintage k and prices that reset are indexed both by the time period they are reset and the vintage of the information used at the point they are reset, $P_{H,t}^k(j,l)$.

The optimality condition for $P_{H,t}^k(j,l)$ is

$$\mathbb{E}_{t-k} \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h} \left(\frac{P_{H,t}^k(j,l)}{P_{t+h}} \right)^{-\theta} c_{t+h}$$
$$= \frac{\theta}{\theta - 1} \mathbb{E}_{t-k} \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h} \left(\frac{W_{t+h}}{P_{H,t}^k(j,l)} \right) \left(\frac{P_{H,t}^k(j,l)}{P_{t+h}} \right)^{-\theta} c_{t+h},$$

(28)

for $k = 0, 1, 2, \cdots$. Canadian firms that sell their good j by exporting to city l also choose prices based on their information set that they last updated. They choose prices so as to solve the maximization problem:

$$\max_{P_{F,t}^k(j,l)} \mathbb{E}_{t-k} \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h}[P_{F,t}^k(j,l) - (1 + \tau(j,l))S_{t+h}W_{t+h}^*]$$
$$\times \left(\frac{P_{F,t}^k(j,l)}{P_{t+h}(j,l)} \right)^{-\theta} \left(\frac{P_{t+h}(j,l)}{P_{t+h}(j)} \right)^{-\theta} \left(\frac{P_{t+h}(j)}{P_{t+h}} \right)^{-\theta} c_{t+h},$$

(29)

for $k = 0, 1, 2, \cdots$. The optimality condition is similar to (28):

$$\mathbb{E}_{t-k} \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h} \left(\frac{P_{F,t}^k(j,l)}{P_{t+h}} \right)^{-\theta} c_{t+h}$$
$$= \frac{\theta}{\theta - 1} \mathbb{E}_{t-k} \sum_{h=0}^{\infty} \lambda_j^h Q_{t,t+h} \left(\frac{(1 + \tau(j,l))S_{t+h}W_{t+h}^*}{P_{F,t}^k(j,l)} \right) \left(\frac{P_{F,t}^k(j,l)}{P_{t+h}} \right)^{-\theta} c_{t+h},$$

(30)

for $k = 0, 1, 2, \cdots$.

2.3.3 Equilibrium

The monetary authority in each country sets the growth rate of the money stock such that it follows an AR(1):

$$\ln \mu_t = \rho \ln \mu_{t-1} + \varepsilon_t, \tag{31}$$

$$\ln \mu_t^* = \rho \ln \mu_{t-1}^* + \varepsilon_t^*, \tag{32}$$

where ε_t and ε_t^* are mean-zero i.i.d shock and $\mu_t = M_t/M_{t-1}$ and $\mu_t^* = M_t^*/M_{t-1}^*$. The steady state (log) money growth rates is set to zero and the common persistence parameter satisfies $\rho \in [0,1)$.

Total transfers from the government in the each country equal domestic money injections minus the lump sum tax from the government paying interest. For the U.S., we have $T_t = M_t - M_{t-1} - (R_{t-1}-1)W_{t-1}n_{t-1}$. The total transfer in Canada is of the same form up to currency conversions: $T_t^* = M_t^* - M_{t-1}^* - (\frac{S_{t-1}R_{t-1}}{S_t} - 1)W_{t-1}^* n_{t-1}^*$.

The profits of U.S. firms accrue exclusively to U.S. households. In other words, $\Pi_t = \int_j \int_{z=0}^{\frac{1}{2}} \Pi_t(j,z)dzdj$, where $\Pi_t(j,z)$ is the profit of a U.S. firm. Similarly, the profits of Canadian firms accrue exclusively to Canadian households: $\Pi_t^* = \int_j \int_{z=\frac{1}{2}}^{1} \Pi_t^*(j,z)dzdj$, where $\Pi_t^*(j,z)$ is the profit of a Canadian firm.

Recall, market clearing conditions for good markets were given by (21) and (22). The labor market clearing conditions are

$$n_t = \int_j \int_{z=0}^{\frac{1}{2}} n_t(j,z)dzdj,$$

$$n_t^* = \int_j \int_{z=\frac{1}{2}}^{1} n_t^*(j,z)dzdj.$$

Last, but not least, the bond market clears at each date: $B_t + B_t^* = 0$ for all t.

An *equilibrium* of the Calvo pricing economy is a collection of allocations and prices:

- $\{c_t(j,l,z)\}_{j,l,z}$, M_t, B_{t+1}, n_t for U.S. households;

- $\{c_t^*(j,l^*,z)\}_{j,l,z}$, M_t^*, B_{t+1}^*, n_t^* for Canadian households;

- $\{P_t(j,l,z), P_t^*(j,l^*,z), n_t(j,z), y_t(j,z)\}_{j,l,z\in[0,1/2]}$ for U.S. firms;

- $\{P_t(j,l,z), P_t^*(j,l^*,z), n_t^*(j,z), y_t^*(j,z)\}_{j,l^*,z\in(1/2,1]}$ for Canadian firms;

- Nominal wages and bond prices satisfy the following conditions:

    1. Households' allocations solve their maximization problem;

2. Prices and allocations of firms solve their maximization problem (23) and (25);

3. All markets clear;

4. The money supply process and transfers satisfy the specifications above.

An equilibrium of the dual stickiness pricing economy is not much different from the definition of the equilibrium of the Calvo pricing economy. Prices and allocations of firms solve the maximization problems (27) and (29) instead of (23) and (25).

3 Model predictions for LOP deviations

We now discuss implications of Kehoe-Midrigan model under Calvo pricing and dual stickiness pricing for the persistence and volatility of deviations from the LOP.

3.1 Calvo pricing

Log-linearization of (24) around the steady state yields the (log) optimal price for U.S. firms that reset prices in period t:

$$\hat{P}_{H,t}(j,l) = (1 - \lambda_j \beta) \sum_{h=0}^{\infty} (\lambda_j \beta)^h \mathbb{E}_t \hat{M}_{t+h}, \tag{33}$$

where $\hat{P}_{H,t}(j,l)$ and $\hat{M}_t$ are the log-deviation of $P_{H,t}(j,l)$ and M_t from the steady state, respectively. Here, we use the proportionality of nominal wages to money supply (i.e., (16)) to replace the log-deviation of W_t with $\hat{M}_t$ (i.e., $\hat{W}_t = \hat{M}_t$). Thus, the firms that adjust prices in period t choose their price to equalize it to the weighted average of the current and future path of nominal marginal costs.

Analogously, we can derive the log-deviation of optimal price for Canadian firms from (26):

$$\hat{P}_{F,t}(j,l) = (1 - \lambda_j \beta) \sum_{h=0}^{\infty} (\lambda_j \beta)^h \mathbb{E}_t (\hat{S}_{t+h} + \hat{M}^*_{t+h}).$$

Substituting out the equilibrium nominal exchange rate, using (19), gives us

$$\hat{P}_{F,t}(j,l) = (1 - \lambda_j \beta) \sum_{h=0}^{\infty} (\lambda_j \beta)^h \mathbb{E}_t \hat{M}_{t+h}. \tag{34}$$

Thus, $\hat{P}_{F,t}(j,l) = \hat{P}_{H,t}(j,l)$, under our specific preference assumption and the log-deviation of price index for $\hat{P}_t(j,l)$ under Calvo pricing becomes

$$\hat{P}_t(j,l) = \lambda_j \hat{P}_{t-1}(j,l) + (1 - \lambda_j) \hat{P}_{H,t}(j,l).$$

It is convenient to normalize $\hat{P}_{H,t}(j,l)$ (and $\hat{P}_t(j,l)$) by $\hat{M}_t$ to assure stationarity. The deviation reset prices from their steady-state relative to the movement in the nominal money supply is

$$\hat{p}_{H,t}(j,l) = (1 - \lambda_j\beta)\sum_{h=0}^{\infty}(\lambda_j\beta)^h\mathbb{E}_t(\hat{M}_{t+h} - \hat{M}_t) = \left[\frac{\lambda_j\beta\rho}{1 - \lambda_j\beta\rho}\right]\hat{\mu}_t, \tag{35}$$

where $\hat{p}_{H,t}(j,l) = \hat{P}_{H,t}(j,l) - \hat{M}_t$ and $\hat{\mu}_t = \hat{M}_t - \hat{M}_{t-1}$. As it turns out $\hat{p}_{F,t}(j,l) = \hat{P}_{F,t}(j,l) - \hat{M}_t = \hat{p}_{H,t}(j,l)$ so the short-run dynamics of the optimal prices are the same for home and foreign firms selling the same good at the same location in spite of the transportation costs which drive a wedge between the prices in the long-run.

The same normalization for the price deviation for good j in city l yields

$$\hat{p}_t(j,l) = \lambda_j\hat{p}_{t-1}(j,l) - \lambda_j\hat{\mu}_t + (1 - \lambda_j)\left[\frac{\lambda_j\beta\rho}{1 - \lambda_j\beta\rho}\right]\hat{\mu}_t, \tag{36}$$

where $\hat{p}_t(j,l) = \hat{P}_t(j,l) - \hat{M}_t$.

The analogous expression for the Canadian price index for good j and city l^* is

$$\hat{p}_t^*(j,l^*) = \lambda_j\hat{p}_{t-1}^*(j,l^*) - \lambda_j\hat{\mu}_t^* + (1 - \lambda_j)\left[\frac{\lambda_j\beta\rho}{1 - \lambda_j\beta\rho}\right]\hat{\mu}_t^*, \tag{37}$$

and the log bilateral real exchange rate for good j across cities l and l^* is $\hat{q}_t(j,l,l^*) = \ln q_t(j,l,l^*) - \ln q(j,l,l^*)$, where $q_t(j,l,l^*)$ is given by

$$q_t(j,l,l^*) = \frac{S_tP_t^*(j,l^*)}{P_t(j,l)}, \tag{38}$$

and $q(j,l,l^*)$ is its steady state value.

The next proposition characterizes the short-run good-level real exchange rate dynamics under Calvo pricing with a slight generalization of Kehoe and Midrigan (2007).

Proposition 1. *Under the preference assumption $U(c,n) = \ln c - \chi n$, the CIA constraints, the assumption of money growth (31) and (32) and good-specific Calvo pricing, the good-level real exchange rate between any cities l and l^* follows an AR(2) process of the form:*

$$\hat{q}_t(j,l,l^*) = (\lambda_j + \rho)\hat{q}_{t-1}(j,l,l^*) - \lambda_j\rho\hat{q}_{t-2}(j,l,l^*) + \theta_j\eta_t, \tag{39}$$

where $\hat{q}_t(j,l,l^*) = \hat{S}_t + \hat{P}_t^*(j,l^*) - \hat{P}_t(j,l)$, $\theta_j = \lambda_j - (1 - \lambda_j)\frac{\lambda_j\beta\rho}{1-\lambda_j\beta\rho}$, and $\eta_t(= \varepsilon_t - \varepsilon_t^*)$ is i.i.d. $(0, \sigma_\eta^2)$.

Proof. From (18) and (19), $\hat{q}_t(j,l,l^*) = \hat{p}_t^*(j,l^*) - \hat{p}_t(j,l)$. Subtracting (36) from (37) yields $\hat{q}_t(j,l,l^*) = \lambda_j\hat{q}_{t-1}(j,l,l^*) + \theta_j(\hat{\mu}_t - \hat{\mu}_t^*)$. Because $\hat{\mu}_t - \hat{\mu}_t^*$ follow an AR(1) from (31) and (32), we obtain (39) and proved Proposition 1. □

Proposition 1 of Kehoe and Midrigan (2007) is a special case of the one above: when money growth rates follow an i.i.d. process ($\rho = 0$) equation (39) reduces to an AR(1) model with its coefficient λ_j and $\theta_j = \lambda_j$ as Kehoe and Midrigan (2007) prove.

3.1.1 Persistence

Turning to implications for persistence of the good-level real exchange rates we employ the sum of autoregressive coefficients (SAR) as the persistence metric. This is often the case in applied work when moving beyond the AR(1) model (e.g., Andrews and Chen (1994) and Clark (2006)) because the SAR has a one-to-one relationship to the cumulative long-run impulse response to a shock. We denote the SAR by α_j.

Under Proposition 1, the SAR measure of persistence is $\alpha_j = \lambda_j + \rho(1 - \lambda_j)$; it simplifies to $\alpha_j = \lambda_j$ when $\rho = 0$. Obviously, SAR is strictly increasing in ρ regardless of the degree of price stickiness under $\lambda_j \in [0, 1)$. The left panel of Fig.1 shows the effect of increasing ρ on the persistence for the two goods: a good with relatively slow price adjustment ($\lambda_j = 0.95$) and a good with relatively fast price adjustment ($\lambda_j = 0.5$).

The right panel of Fig.1 plots the SAR against λ_j. The figure compares the model's implications for $\rho = 0$, as calibrated by Kehoe and Midrigan (2007) and $\rho = 0.83$, the monthly analog to the CKM calibration.[6] The impact of introducing persistence in money growth rates on the SAR is clear. When $\rho = 0$, the model predicts that the SAR equals λ_j, so the two lie on the 45 degree line in the figure. On the other hand, when $\rho > 0$, the model predicts a much flatter line. Thus, a high persistence of the money growth rates increases the persistence of LOP deviations, regardless of the frequency of price adjustment, but the quantitative impact is greatest when the frequency of price adjustment is highest.

To see the intuition behind the persistent dynamics it is instructive to express the current LOP deviation as a function of its lagged self and the change in the nominal exchange rate:

$$\hat{q}_t(j, l, l^*) = \lambda_j \hat{q}_{t-1}(j, l, l^*) + \theta_j \Delta \hat{S}_t, \tag{40}$$

where $\Delta \hat{S}_t = \hat{\mu}_t - \hat{\mu}_t^*$ from (19). When $\rho = 0$ as in Kehoe and Midrigan (2007), $\Delta \hat{S}_t$ is an i.i.d shock and the good-level real exchange rate follows AR(1) with persistence parameter, λ_j. When

---

[6] The CKM estimate of the autoregressive coefficient is 0.68 using quarterly U.S. data for M1 growth. We transform this quarterly persistence of M1 growth into the monthly persistence by solving $Cov(\hat{M}_t - \hat{M}_{t-3}, \hat{M}_{t-3} - \hat{M}_{t-6})/Var(\hat{M}_t - \hat{M}_{t-3}) = 0.68$ for ρ. We obtained the resulting monthly persistence of M1 money growth of 0.83.

international money growth differential is positively autocorrelated ($\rho > 0$) so is the change in the nominal exchange rate, which contributes to increased persistence in the real exchange rate.

3.1.2 Volatility

Throughout, real exchange rate volatility will be measured relative to the standard deviation of the change in the nominal exchange rate: $\sigma_j = \text{std}(q_t(j, l, l^*))/\text{std}(\Delta S_t)$. When $\rho = 0$, the model predicts the normalized standard deviation to be $\sigma_j = \sigma_1(\lambda_j) = \lambda_j/\sqrt{1 - \lambda_j^2}$ and a good with larger λ_j will exhibit more variability. When $\rho > 0$, the normalized standard deviation is predicted to be of the form $\sigma_j = \sigma_2(\lambda_j, \rho, \beta)$ and may be obtained using the variance formula of an AR(2) process along with $\text{std}(\Delta S_t) = \text{std}(\eta_t)/\sqrt{1 - \rho^2}$. Importantly, the volatility function depends not just on λ_j, but also on ρ and β.

An implication of this is that increased persistence in money growth, while helpful in resolving the persistence puzzle, may actually make the volatility puzzle worse because $\sigma_j = \sigma_2(\lambda_j, \rho, \beta)$ turns out not to be monotonic in ρ. Even more disturbing is that the shape of the relationship with ρ depends on the frequency of price adjustment, which we know differs across goods. The practical thrust of this is: changes in money growth persistence will have differential impacts across goods.

The left panel of Fig.2 plots the normalized standard deviations $\sigma_j = \sigma_2(\lambda_j, \rho, \beta)$ against ρ.[7] For a good with relatively infrequent price changes ($\lambda_j = 0.95$), volatility of the real exchange rate rises over most of the range of money growth persistence, before falling sharply as money growth approaches a random walk. In contrast, for a good with relatively frequent price changes ($\lambda_j = 0.5$), the volatility of the relative price is declining in the money growth rate throughout. The right panel of Fig.2 shows the ambiguous impact of introducing a positive ρ on the volatility from another dimension. The normalized standard deviation is smaller for $\rho = 0.83$ than for $\rho = 0$ when price adjustment is fast. When the price adjustment is slow, we have a larger normalized standard deviation for $\rho = 0.83$ than for $\rho = 0$.

3.2 Calvo pricing with infrequent information updating

Let $\hat{P}_{H,t}^k(j, l)$ be the log deviation of $P_{H,t}^k(j, l)$ from the steady state. Log-linearizing (28) around the steady state yields

$$\hat{P}_{H,t}^k(j, l) = (1 - \lambda_j \beta) \sum_{h=0}^{\infty} (\lambda_j \beta)^h \mathbb{E}_{t-k} \hat{M}_{t+h}, \text{ for } k = 0, 1, 2, \cdots.$$

---

[7]We set the discount factor β to 0.99.

The law of iterated expectations implies

$$\hat{P}_{H,t}^k(j,l) = \mathbb{E}_{t-k}\hat{P}_{H,t}(j,l).$$

Here, we use $\hat{P}_{H,t}^0(j,l) = \hat{P}_{H,t}(j,l)$ because of the equivalence between (24) and (28) when $k = 0$.

Consider the weighted average of newly set prices that U.S. firms choose when they adjust prices in period t; these firms choose $\mathbb{E}_{t-k}\hat{P}_{H,t}(j,l)$ according to their information they last updated. Canadian firms choose $\mathbb{E}_{t-k}\hat{P}_{F,t}(j,l)$. As before $\hat{P}_{F,t}(j,l) = \hat{P}_{H,t}(j,l)$. Therefore, $\hat{P}_{F,t}^k(j,l) = \hat{P}_{H,t}^k(j,l)$ for $k > 0$, due to the law of iterated expectations.

Defining $\hat{X}_t(j,l)$ as the weighted average for the newly set prices for good j in city l of the U.S., based upon different information vintages, we obtain

$$\hat{X}_t(j,l) = (1 - \omega_j)\sum_{k=0}^{\infty}\omega_j^k\mathbb{E}_{t-k}\hat{P}_{H,t}(j,l), \tag{41}$$

which is similar in mathematical formulation to the price index in Mankiw and Reis (2002, p.1300).

Now, using the definition $\hat{P}_{H,t}(j,l) = \Delta\hat{P}_{H,t}(j,l) + \hat{P}_{H,t-1}(j,l)$, (41) can be rewritten as

$$\hat{X}_t(j,l) = (1 - \omega_j)\hat{P}_{H,t}(j,l) + \omega_j(1 - \omega_j)\sum_{k=0}^{\infty}\omega_j^k\mathbb{E}_{t-k-1}\Delta\hat{P}_{H,t}(j,l)$$

$$+ \omega_j(1 - \omega_j)\sum_{k=0}^{\infty}\omega_j^k\mathbb{E}_{t-k-1}\hat{P}_{H,t-1}(j,l).$$

The second line of the equation is $\omega_j\hat{X}_{t-1}(j,l)$ from (41). Hence,

$$\hat{X}_t(j,l) = \omega_j\hat{X}_{t-1}(j,l) + (1 - \omega_j)\hat{P}_{H,t}(j,l) + \omega_j(1 - \omega_j)\sum_{k=0}^{\infty}\omega_j^k\mathbb{E}_{t-k-1}\Delta\hat{P}_{H,t}(j,l).$$

To render the variable stationary, define $\hat{x}_t(j,l) = \hat{X}_t(j,l) - \hat{M}_t$. Then,

$$\hat{x}_t(j,l) = \omega_j\hat{x}_{t-1}(j,l) - \omega_j\hat{\mu}_t + (1 - \omega_j)\hat{p}_{H,t}(j,l)$$

$$+ \omega_j(1 - \omega_j)\sum_{k=0}^{\infty}\omega_j^k\mathbb{E}_{t-k-1}[\Delta\hat{p}_{H,t}(j,l) + \hat{\mu}_t]. \tag{42}$$

Appendix B shows that we can derive the closed form solution to $\hat{x}_t(j,l)$:

$$\hat{x}_t(j,l) = \omega_j\hat{x}_{t-1}(j,l) + a_j\hat{\mu}_t + \frac{b_j}{1 - \omega_j\rho L}\hat{\mu}_{t-1}, \tag{43}$$

where $a_j = \frac{\lambda_j\beta\rho - \omega_j}{1 - \lambda_j\beta\rho}$, $b_j = \frac{\omega_j\rho(1 - \lambda_j\beta)(1 - \omega_j)}{1 - \lambda_j\beta\rho}$, and L is the lag operator.

The price index for good j in city l is a Calvo-weighted-average of fixed and reset prices. The latter being our weighted average of price resets given different vintages of information:

$$\hat{P}_t(j,l) = \lambda_j\hat{P}_{t-1}(j,l) + (1 - \lambda_j)\hat{X}_t(j,l).$$

16

Again, normalizing by $\hat{M}_t$, gives

$$\hat{p}_t(j,l) = \lambda_j \hat{p}_{t-1}(j,l) - \lambda_j \hat{\mu}_t + (1-\lambda_j)\hat{x}_t(j,l) \tag{44}$$

and Canadian versions of these expressions are:

$$\hat{x}_t^*(j,l^*) = \omega_j \hat{x}_{t-1}^*(j,l^*) + a_j \hat{\mu}_t^* + \frac{b_j}{1-\omega_j \rho L}\hat{\mu}_{t-1}^*,$$
$$\hat{p}_t^*(j,l^*) = \lambda_j \hat{p}_{t-1}^*(j,l) - \lambda_j \hat{\mu}_t^* + (1-\lambda_j)\hat{x}_t^*(j,l^*).$$

The next proposition establishes the rich short-run dynamics of the good-level real exchange rate emerging from the extended model.

Proposition 2. *Under the preference assumption* $U(c,n) = \ln c - \chi n$, *the CIA constraints, the assumption of money growth (31) and (32), along with good-specific Calvo pricing and good-specific Mankiw-Reis information updating, the good-level real exchange rate between any cities l and l^* follows an ARMA(4,2) process of the form:*

$$\hat{q}_t(j,l,l^*) = \sum_{r=1}^{4} \phi_{j,r}\hat{q}_{t-r}(j,l,l^*) + \sum_{r=0}^{2} \theta_{j,r}\eta_{t-r} \tag{45}$$

where

$$\phi_{j,1} = \tilde{\phi}_{j,1} + \rho, \qquad \tilde{\phi}_{j,1} = \lambda_j + \omega_j + \omega_j \rho$$
$$\phi_{j,2} = \tilde{\phi}_{j,2} - \tilde{\phi}_{j,1}\rho, \qquad \tilde{\phi}_{j,2} = -[\lambda_j \omega_j + (\lambda + \omega_j)\omega_j \rho]$$
$$\phi_{j,3} = \tilde{\phi}_{j,3} - \tilde{\phi}_{j,2}\rho, \qquad \tilde{\phi}_{j,3} = \lambda_j \omega_j^2 \rho$$
$$\phi_{j,4} = -\tilde{\phi}_{j,3}\rho$$
$$\theta_{j,0} = \lambda_j - (1-\lambda_j)a_j$$
$$\theta_{j,1} = -\lambda_j(\omega_j + \omega_j \rho) + (1-\lambda_j)(\omega_j \rho a_j - b_j)$$
$$\theta_{j,2} = \lambda_j \omega_j^2 \rho.$$

Proof. See Appendix C. □

When $\omega_j = 0$ this proposition reduces to Proposition 1.[8] Below, we discuss that both the persistence and volatility of good-level real exchange rates predicted by the dual stickiness pricing can be quite high. Moreover, this is true even if the price adjustment is relatively fast, which is essential in matching the cross-sectional evidence which contains goods with frequent price changes and, yet, high persistence and variability in their LOP deviations.

---

[8]In particular, we obtain $\phi_{j,1} = \lambda_j + \rho$, $\phi_{j,2} = -\lambda_j \rho$, and $\phi_{j,3} = \phi_{j,4} = 0$ for the AR parameters and $\theta_{j,0} = \theta_j$ and $\theta_{j,1} = \theta_{j,2} = 0$ for the MA parameters.

3.2.1 Persistence

Appendix C shows that the SAR in this generalized case is given by

$$\alpha_j = \sum_{r=1}^{4} \phi_{j,r} = 1 - (1 - \lambda_j)(1 - \omega_j)(1 - \omega_j \rho)(1 - \rho).$$

Clearly, the slower the speed of information updating adjustment is ($\omega_j \to 1$), the larger the SAR becomes.

For a general ARMA process without parameter restrictions, it is not conventional to use the SAR as a measure of persistence, because of the presence of MA terms. However, if our model is correctly specified, we can show that both the long-run impact of cumulative impulse response of a unit monetary shock on real exchange rates and the SAR are strictly increasing function of λ_j, ω_j, and ρ. Furthermore, using the SAR is also convenient in computation and for the purpose of making comparison with simpler models introduced in the previous subsection. For these reasons, we continue to focus on the SAR as an approximate measure of persistence under the assumption that the process (45) is correctly specified.

The extended model works well in generating the persistence of a good-level real exchange rate. The left panel of Fig.3 shows the SAR among different ω_j's. The persistence is increasing in ω_j and is very high regardless of the infrequency of price changes.[9] The right panel of Fig.3 plots the persistence against λ_j. This panel compares cases of two extreme values of ω_j. One is the case in which firms producing good j updates their information every month. (i.e., $\omega_j = 0$.) The other is the case in which firms, on average, update information every 50 months (i.e., $\omega_j = 0.98$). For the former case, the obtained SAR corresponds to the upper straight line in the right panel of Fig.1 since we set $\rho = 0.83$ in the computation. In the latter case, the persistence measure is very close to one whether prices are sticky or flexible.

3.2.2 Volatility

Having improved the potential of the model in accounting for persistence of real exchange rates, we ask if it helps along the dimension that was more ambiguous in the baseline model, variability. We calculate the new normalized standard deviation $\sigma_j = \sigma_3(\omega_j, \lambda_j, \rho, \beta)$, using the fact that the good-level real exchange rates now follow the ARMA(4,2) process according to Proposition 2. The left panel of Fig.4 plots the normalized standard deviations against ω_j. It shows that the volatility

---

[9]Even if $\omega_j = 0$, $\hat{q}_t(j, l, l^*)$ has been already somewhat persistent, because of the AR(1) money growth.

grows exponentially as ω_j increases. The right panel of Fig.4 shows the effect of increasing λ_j on the normalized standard deviations under the two extreme cases: $\omega_j = 0$ and 0.98. It shows that real exchange rate volatility becomes substantially greater when the information adjustment is slower. Thus, the introduction of information stickiness enhances the real exchange rate volatility to a large extent.

The question we pose next is what lengths of information delays do we need to match key properties of the micro-data, conditional on the model. The key properties are the persistence and volatility of good-level real exchange rates along with the frequency of price changes observed for those same goods.

4 Empirical results

4.1 Data

The source of our retail prices is the *Worldwide Cost of Living Survey* compiled by the Economist Intelligence Unit (EIU). It is an extensive annual survey of international retail prices that was originally designed to help managers to determine compensation levels of their employees residing in different cities of the world. The coverage of goods and services is broad enough to overlap significantly with what appears in a typical urban consumption basket (see Rogers (2007), for more detail on the comparison between EIU data and the CPI data from national statistical agencies). A notable advantage of the EIU data is the fact that all the individual good prices are listed in absolute terms with the survey conducted by a single agency in a consistent manner over time. Because of this convenient panel data format, a number of recent studies on international price dynamics have used this data, including Crucini and Telmer (2007), Crucini and Shintani (2008), Engel and Rogers (2004), Parsley and Wei (2007) and Rogers (2007).

For a limited number of countries, the EIU data contains observations from multiple cities. In our empirical analysis, we focus on U.S.-Canadian city pairs since the assumption of the common probability of price adjustment for each good seems to be a reasonable approximation between the two neighboring countries.[10] After removing missing observations to construct a balanced panel for the period from 1990 to 2005, 3 of the 16 available U.S. cities available in the survey are dropped, while all 4 Canadian cities remain. This results in a total of 52 unique city pairs. The cities and

---

[10]Alternatively, one may use the average of price change frequencies between the two countries, an approach employed in Kehoe and Midrigan (2007), when data from both countries are available.

categories of goods included in the analysis are shown in Fig. 5 and Table A1, respectively.

For each good j, the log of $q_t(j, l, l^*)$ for each year t $(= 1, ..., 16)$ is computed using the price level in a U.S. city l $(= 1, ..., 13)$ expressed in U.S. dollars $(P_t(j, l))$, the price level in a Canadian city l^* $(= 1, ..., 4)$ expressed in Canadian dollars $(P_t^*(j, l^*))$, and the spot U.S.–Canadian dollar exchange rate (S_t), all from the EIU data. Since the resulting log real exchange rates represent the log deviations of the price in a Canadian city relative to that of a U.S. city both expressed in a common currency, a negative value for the pair of Toronto and New York, for example, implies that the good is more expensive in New York than in Toronto at year t. Fig.6 plots the log of $q_t(j, l, l^*)$, pooling all goods and all city pairs from two selected years, 1990 and 2005.

Next, for the price stickiness parameter, λ_j, we utilize the frequency of price changes, f_j and transform it with $\lambda_j = 1 - f_j$ for good j. Since the EIU data is annual, it is not useful for constructing monthly frequency of price changes. For this part of our analysis, we rely on existing studies based on monthly micro-data from the BLS (Bureau of Labor Statistics). Bils and Klenow (2004) used the BLS Commodities and Services Substitution Rate Table for 1995-1997 which contains the average frequencies of price changes of individual goods and services used in construction of the U.S. CPI. We took the monthly average frequency of price changes, f_j, from Table A1 of their paper and matched them with the 165 goods in the EIU sample. Since we require persistence and frequency adjustment parameters good-by-good to evaluate the model, we use only these 165 matched pairs in our analysis. We assume that the frequency of price changes applies to the entire sample period of 1990-2005 in our EIU data set.[11] In addition, we assume a common frequency of price change between the U.S. and Canadian cities, good-by-good.

Nakamura and Steinsson (2007) recently revisited Bils and Klenow's analysis using more detailed and updated BLS data. Using the CPI Research Database created by the BLS, they re-estimated the frequencies of price change after removing temporary price changes associated with sales. They found that the median duration between regular price changes was 8 - 11 months depending on the treatment of substitutions, considerably higher than the 4.3 months for the median good, found by Bils and Klenow (2004). In what follows, we also check the impact on our results of using the Nakamura and Steinsson's (2007) data on the frequency of price changes from the period of

---

[11]In some countries which experienced a structural shift in inflation, an assumption of constant frequency of price changes over years may not be satisfied. For example, Ahlin and Shintani (2007) use Mexican price data on 44 goods and report that the average monthly frequency of price changes was 28% in 1994 and as large as 50% in 1995. We expect that this issue is less serious in our case since both U.S. and Canada had a stable inflation during the period under consideration.

1998-2005.

For the nominal exchange rate changes required for the theoretical volatility calculation, we use monthly changes in the log of the end-of-month U.S.-Canadian dollar spot rates. While both price stickiness parameter (frequency of price changes) and nominal exchange rates are available in monthly series, real exchange rates are only observed annually. The small number of time series observation at the annual frequency is the major limitation of the EIU data. In the next subsection, we briefly discuss how to reconcile the mixed frequencies of observation in the dynamic panel estimation and describe the procedure to estimate the time series models.

4.2 Estimation

Table 1 shows how monthly ARMA processes predicted by the model are transformed into the ones which have non-zero coefficients for multiples of 12 month lags and finite MA terms. The first row of the table shows the easiest transformation. In Calvo pricing with $\rho = 0$, the equation (39) directly implies that

$$\hat{q}_t(j,l,l^*) = \lambda_j \hat{q}_{t-1}(j,l,l^*) + \lambda_j \eta_t.$$

By repeated substitutions, we get

$$\hat{q}_t(j,l,l^*) = \lambda_j^{12} \hat{q}_{t-12}(j,l,l^*) + \lambda_j \Lambda_j(L) \eta_t$$

where $\Lambda_j(L) = \sum_{r=0}^{11} \lambda_j^r L^r$. In this equation, the AR term is the 12th lag (in months) and the order of the MA term is 11. This ARMA(12,11) is equivalent to an AR(1) sampled annually since $\lambda_j \Lambda_j(L) \eta_t$ and $\hat{q}_{t-12}(j,l,l^*)$ are not correlated.

Such a transformation is not necessarily possible with a general ARMA process including AR(2) and ARMA(4,2) processes. However, thanks to a special dynamic feature of the theoretical model, it is possible that we can make the AR parameters non-zero only if the lags are multiples of 12 and the MA parameters finite under our extended models (39) and (45). Appendix D provides the detailed derivations of these more elaborate transformations.

Previously, l and l^* were used for the U.S. and Canadian cities, respectively. Here, they are replaced by a new single index i ($= 1, ..., 52$) each representing a city pair spanning a national border. In addition, the sampling frequency for the model was assumed to be monthly. With some abuse of notation, our new time subscript now represents the time in annual frequency. Namely, if the true data process is generated for each month $t^* = 1, ..., T^*$, we now only observe the series

annually at the months of $t = 12 \times t^* = 1, ..., T(= T^*/12)$. With this newly introduced index, we define q_{it}^j as the *log* of the real exchange rate for good j between the city pair i at year t:

$$q_{it}^j = \ln q_t(j, l, l^*).$$

Thus, the former log deviation from the steady state $\widehat{q}_t(j, l, l^*)$ can be rewritten as $q_{it}^j - q_i^j$, where q_i^j is the long-run value which the Appendix E derives:

$$q_i^j = \ln q(j, l, l^*) = \ln \frac{[1 + \kappa^{1-\theta}(1 + \tau(j, l^*))^{1-\theta}]^{\frac{1}{1-\theta}}}{[1 + \kappa^{1-\theta}(1 + \tau(j, l))^{1-\theta}]^{\frac{1}{1-\theta}}}.$$

Intuitively, the relative price of a good in the long-run is higher in the destination market with the higher shipping cost from the source. Thus if city l^* is, say, farther from the source of the good than city l, q_i^j is positive. These heterogeneous long-run deviations justify the presence of the individual effect (the time invariant city pair-specific effect) in the panel estimation.

Based on the annual transformation shown in Table 1, all the dynamics of the real exchange rate for good j can be written as

$$q_{it}^j = \sum_{r=1}^m \Phi_{j,r} q_{i,t-r}^j + \zeta_i^j + u_t^j + v_{it}^j,$$

where ζ_i^j is the time invariant unobserved city pair-specific effect which allows long-run price difference between two cities, u_t^j is the common time effect which represents the exchange rate shocks and v_{it}^j is a good-specific residual term.

This model format nests all the models under consideration: (i) Calvo pricing with $\rho = 0$ implies $m = 1$; (ii) Calvo pricing with $\rho \neq 0$ implies $m = 2$; and (iii) dual stickiness pricing implies $m = 4$. For the individual specific effect ζ_i^j, we can easily see its relationship to the long-run mean and the persistence from $q_i^j = \zeta_i^j/(1 - \alpha_j)$ where $\alpha_j = \sum_{r=1}^m \Phi_{j,r}$. For the common time effect u_t^j, Calvo pricing with $\rho \neq 0$ predicts a serial correlation of order one, while dual stickiness pricing predicts a serial correlation of order three. However, in a short panel asymptotic with finite T, the common time effects can be treated as unknown parameters to be estimated with time dummies. In addition, since our main interest is to estimate the persistence expressed in terms of the SAR, α_j, it is convenient to rewrite the model into the augmented Dickey-Fuller (ADF) form. Thus, the nested model is given by

$$q_{it}^j = \alpha_j q_{i,t-1}^j + \sum_{r=1}^{m-1} \gamma_{j,r} \Delta q_{i,t-r}^j + u_j^\top \widetilde{D}_t + \zeta_i^j + v_{it}^j,$$

where $\Delta q_{i,t-r}^j = q_{i,t-r}^j - q_{i,t-r-1}^j$, $\gamma_{j,r} = \sum_{s=r+1}^{m} \Phi_{j,s}$ for $r = 1, ..., k-1$, $u_j = (u_{m+1}^j, ..., u_T^j)^\top$ is a vector of constants, $\widetilde{D}_t$ is a $(T-m) \times 1$ time dummy vector with one in the t-th position and zero otherwise.

To estimate this short dynamic panel model, we employ the generalized method of moments (GMM) estimator in the first differenced form for the purpose of eliminating the individual effect ζ_i^j. We follow Arellano and Bond (1991) in the choice of instruments and initial weighting matrix. In particular, the moment condition is given by

$$\mathbb{E}\left[q_{is}^j\left(\Delta q_{it}^j - \alpha_j \Delta q_{i,t-1}^j - \sum_{r=1}^{m-1} \gamma_{j,r}\Delta^2 q_{i,t-r}^j - \delta_j^\top D_t\right)\right] = 0$$

for $s = 1, ..., t-m-1$ and $t = m+2, ..., T$, where $\Delta^2 q_{i,t-r}^j = \Delta q_{i,t-r}^j - \Delta q_{i,t-r-1}^j$ $\delta_j = (\Delta u_{m+2}^j, ..., \Delta u_T^j)^\top$ is a vector of constants, D_t is a $(T-m-1) \times 1$ time dummy vector with one in the t-th position and zero otherwise. The total number of parameters to be estimated is $T-1$ with the number of moment conditions given by $(T-m)(T-m-1)/2$.[12] This GMM estimator for α_j is consistent under large N fixed T asymptotics.

4.3 Persistence

In this subsection, we evaluate the Kehoe-Midrigan model and its extension in explaining the observed persistence of the real exchange rate for each good j. Following the theoretical analysis, our empirical persistence measure is the SAR α_j.

We first revisit the original Kehoe-Midrigan model with an assumption of an i.i.d. money growth ($\rho = 0$). In this case, the theory predicts an AR(1) model and thus α_j is simply an AR(1) coefficient. A GMM estimation of α_j yields a median of 0.56 using annual U.S.-Canadian city pairs data.[13] In terms of monthly frequency, our value corresponds to $0.56^{1/12} = 0.95$, which is slightly less than 0.98, the median value obtained by Kehoe and Midrigan (2007) based on bilateral real exchange rates of 66 goods between the U.S. and four European countries, Austria, Belgium, France and Spain.

The first panel in Fig.7 plots the estimated persistence measure α_j against the infrequency of the price adjustment in the annual frequency $\lambda_j^{12} = (1 - f_j)^{12}$ computed based on f_j from Bils and

---

[12]For the model to be (over-) identified, at least $T = 4$ is required for $m = 1$, $T = 6$ is required for $m = 2$, and $T = 9$ is required for $m = 4$. Since $T = 16$ is available in our sample, the number of over-identifying restrictions is 51, 76, and 90, respectively, for $m = 1, 2$, and 4.

[13]This value lies between the medians for OECD city pairs (0.65) and LDC city pairs (0.51) obtained by Crucini and Shintani (2008) based on the same data source.

Klenow's (2004) table. A cross-sectional regression of α_j on λ_j^{12} yielded a significantly positive slope coefficient estimate of 0.30 (with a standard error of 0.08) which is consistent with the theoretical prediction at least in direction: more price stickiness implies higher persistence. However, 160 out of 165 goods lie above the 45 degree line ($\alpha_j = \lambda_j^{12}$) in the scatter plot with the regression slope being significantly less than unity. If the model performance is evaluated by computing the ratio of the predicted persistence (on the 45 degree line) to the observed persistence for each good, the model can explain merely a 6 percent of the total persistence for the median good. This confirms Kehoe and Midrigan's claim that a simple model of price stickiness alone is quantitatively insufficient to reproduce the observed persistence in good-level real exchange rates.

We next consider the effect of introducing serially correlated money growth ($\rho = 0.83$). On the whole, the persistence estimate α_j remains almost unchanged with a median value of 0.57 based on the AR(2) model. The regression slope shown in the second panel of Fig.7 is 0.35 and is again significantly positive. Recall that for a given λ_j, α_j is a monotonically increasing function of ρ (see the left panel of Fig.1). In annual frequency, the predicted SAR from Table 1 is given by

$$\alpha_j = 1 - (1 - \rho^{12})(1 - \lambda_j^{12}) = \lambda_j^{12} + \rho^{12} - \lambda_j^{12}\rho^{12}$$

and the effect of increasing ρ can be seen in the median value of the ratio of prediction and data provided in the upper panel of Table 2. In terms of the median, the theoretical persistence becomes the observed persistence when ρ is around 0.95. However, this value is much higher than $\rho = 0.83$, the reference value based on CKM. Indeed, when $\rho = 0.83$ is used, only 31 percent of the persistence can be explained by the model (the number is provided as the first entry of the lower panel). This fact of an insufficient persistence of the money growth in explaining the persistence of real exchange rates can be also seen from the scatter plot. Recall that, from the right panel of Fig.1, increasing ρ shifts the theoretical line upward with a flatter slope. A similar theoretical prediction line with $\rho = 0.83$, expressed in the annual frequency basis, is also drawn in the second panel of Fig.7.[14] Compared to the 45 degree line in the first panel of the same figure ($\rho = 0$), the predicted line now becomes flatter but is still much steeper than the regression line. Indeed, about 95 percent of data points are still above the $\rho = 0.83$ line. Thus, persistence in money growth helps a bit, but the model with Calvo pricing remains largely unsuccessful in explaining the persistence with a reasonable choice of money growth process.

Third, we now look at the role of information delay in explaining α_j. To simplify the argument,

---

[14]The intercept of the theoretical line is $\rho^{12} = 0.83^{12} = 0.11$.

here we assume the information delay parameter to be common across all the goods (namely, $\omega_j = \omega$ for all j). The persistence estimates based on the AR(4) model become somewhat lower with a median value of 0.51, but still are much higher than the level predicted by the standard Calvo pricing without information delay (which corresponds to the $\omega = 0$ line shown in the lower panel of Fig.7). Recall that from the left panel of Fig.3, for a fixed value of λ_j and $\rho(= 0.83)$, α_j is strictly increasing in ω. This pattern is preserved in the SAR expressed in annual frequency (See Appendix D.):

$$\alpha_j = 1 - (1 - \rho^{12})(1 - \lambda_j^{12})(1 - \omega^{12})(1 - (\omega\rho)^{12}).$$

Based on this relationship, median of the ratio of theoretical value to observed value, provided in the lower panel of Table 2, increases with ω and reaches one at $\omega = 0.93$ which corresponds to 14 months of average duration between information updates. Therefore, at least in terms of the median, dual stickiness pricing with a reasonable money growth process is capable of replicating the observed persistence. In the lower panel of Fig.7, the shaded triangle area shows the range between the line without information delay ($\omega = 0$) and the line with an enormous information delay ($\omega = 0.98$ which corresponds to the 50 month average duration of information updates). The regression line is located almost in the middle of the triangle with a slope of 0.56 which lies strictly between the slopes of the upper and lower bound prediction lines.

We now turn to the results based on f_j from Nakamura and Steinsson's (2007) data. Fig.8 shows the scatter plots of the pairs of $(\alpha_j, \lambda_j^{12})$ for (i) Calvo pricing with $\rho = 0$; (ii) Calvo pricing with $\rho > 0$; and (iii) dual stickiness pricing, respectively.

For many goods, the removal of sales results in the lower value of f_j. Less frequent price changes increase the value of $\lambda_j^{12} = (1 - f_j)^{12}$, and most of the data points in the scatter plot shift toward right.[15] For all the models, the predicted persistence will be higher for the larger values of λ_j^{12}, and thus excluding the sales from frequency of price changes works in favor of the Kehoe-Midrigan model's prediction about real exchange rate persistence. The proportion of the data points lie below the theoretical prediction line increased from 3 percent to 16 percent for Calvo pricing with $\rho = 0$, and from 5 percent to 24 percent for Calvo pricing with $\rho = 0.83$. For all the case, regression slopes shown in the scatter plots are again significant and positive, and the regression fit in terms of the coefficient of determination becomes uniformly better.[16] However, because of the rightward shift,

---

[15]Note that α_j for each good j for each AR model remains unchanged between Figs.7 and 8. In addition, because the sample periods differ between the two data sets, this rightward shift may not be true for some goods.

[16]Regression coefficients are 0.36, 0.31, and 0.43 for each panel, respectively. Coefficients of determination increase

more data points in the last panel of the figure fall outside the shaded triangle region representing the theoretical predictions of dual stickiness pricing.

We see improvement in the ratio of predicted to the observed persistence provided in Table 3. As shown in the first entry of the upper panel of Table 3, even the case with $\rho = 0$ can account for 48 percent of the observed persistence, in comparison to 6 percent based on Bils and Klenow's f_j. The ratio increases as ρ increases, but because of the higher initial ratio, it becomes one at around $\rho = 0.92$ a value lower than previously selected value of $\rho = 0.95$. This newly selected ρ, however, is again higher than the CKM's reference value of $\rho = 0.83$. Since the ratio is 66 percent at $\rho = 0.83$, there is still a room for the information delay structure to fill the gap between the theoretical and observed value. The lower panel of Table 3 shows the effect of increasing ω on the prediction ratio based on Nakamura and Steinsson's data. The table shows that the 100 percent of the persistence can be explained at about $\omega = 0.90$ which corresponds to 9.5 months of average duration between information updates. This length of months suggests that the role of information stickiness remains important even when Bils and Klenow estimates are replaced with Nakamura and Steinsson estimates.

4.4   Volatility

The second puzzle brought up in Kehoe and Midrigan's (2007) study is the observation of too much volatility in good-level real exchange rates which can neither be explained by a simple sticky price model nor a model with pricing complementarities. In this subsection, we evaluate the role of information stickiness in terms of explaining the observed volatility.

The performance of the model is evaluated by the ratio of the 'theoretical' normalized standard deviation to the 'observed' normalized standard deviation. The procedure of computing each standard deviation is as follows. First, to compute 'theoretical' normalized standard deviation, note that the standard deviation of real exchange rates predicted by the theory has the same implication to both annually sampled data and monthly sampled data. Therefore, unlike the measure of persistence that required transformation shown in Table 1, using annual data requires no further complication. For each good, the theoretical normalized standard deviation σ_j can be directly obtained by substituting $\lambda_j = 1 - f_j$ into the formula discussed in Section 3.

Second, to compute the 'observed' normalized standard deviation, note that using a pooled sample variance as a volatility measure is not appropriate since it includes the variance component

---

from 8 to 26 percent, 10 to 17 percent, and 12 to 15 percent, respectively.

due to the dispersion of long-run real exchange rate q_i^j among city pairs in our panel data. In addition, the theory predicts volatility caused by the nominal exchange rate fluctuation which is common to all the products, but is not designed to incorporate the idiosyncratic variance component such as the one due to time-varying city specific shocks. For this reason, we conduct a variance decomposition based on a standard two-way error components model and focus on the extracted variance component due to a time specific shock. This decomposition seems to be a reasonable choice in our study because it is consistent with the idea of using time dummies in the dynamic panel estimation to incorporate the common time specific shocks in our previous analysis of persistence. We thus use the observed standard deviation of time specific component normalized by the sample standard deviation of monthly nominal exchange rate growth.

We start with looking at the results presented in Table 4 based on Bils and Klenow's data. The upper panel of the table shows the median of the ratio of the theoretical to observed normalized standard deviation. The original Kehoe-Midrigan setting with $\rho = 0$ can explain only 13 percent of the variation in the data. Thus, the evidence of excess volatility discovered by Kehoe and Midrigan (2007) is also confirmed in our panel data of the U.S.-Canadian city pairs. Can we explain this observed volatility with an introduction of serially correlated money growth? Unfortunately, unlike the persistence, the predicted volatility is not a monotonically increasing function of ρ. Examples presented in the left panel of Fig.2 show that the volatility decreases monotonically for goods with small $\lambda_j = 1 - f_j$ and increases only in some range of ρ for goods with a larger λ_j. As a result of the combination of the two effects for many goods, none of the median ratio presented in the upper panel of Table 4 is above one and maximum value is only 15 percent at $\rho = 0.52$.

In contrast to the effect of ρ, the left panel of Fig.4 shows that the volatility increases monotonically with ω in dual stickiness pricing for any given values of λ_j and ρ. The lower panel of Table 4 presents the ratio of standard deviations based on dual stickiness pricing with various ω's when the CKM's reference value of $\rho = 0.83$ is used for the money growth process. With an introduction of the information delay, the volatility can now be fully explained at $\omega = 0.94$ which implies 17 months of average duration between information updates. As shown in the previous section, the observed persistence can be reproduced if any value of ρ is allowed without introducing information stickiness. For the volatility, however, the observation can be replicated only under dual stickiness pricing. In this sense, the information delay plays an essential role in explaining volatility.

We now turn to Nakamura and Steinsson's data with the effect of sales removed from f_j. The median of the ratio of the predicted to observed standard deviation for each pricing is shown in

Table 5. The performance of the Kehoe-Midrigan model, in terms of explaining volatility, clearly improves over the results in Table 4 based on Bils and Klenow's data. Since the removal of sales results in the lower values of f_j, using Nakamura and Steinsson's data increases the theoretical volatility level which is increasing in $\lambda_j = 1 - f_j$. For example, when $\rho = 0$, the ratio increases from 13 to 23 percent. However, the degree of increased theoretical volatility is still insufficient to fully explain the observed volatility under the model without information delay. As shown in the upper panel of Table 5, the maximum ratio is only 43 percent under the model without information delay at $\rho = 0.80$. Therefore, the role of sticky information is again crucial in explaining the volatility of the good-level real exchange rates, even if we use Nakamura and Steinsson's data. The lower panel of Table 5 shows that the model with information delay can explain 100 percent of observed volatility when $\omega = 0.92$ which implies 12 months of average duration between information updates. In comparison to the result from Bils and Klenow's data, the reduction of ω reflects the fact that a larger component in the variance is already explained by the reduction of price change frequency alone in Nakamura and Steinsson's data.

4.5 Good-specific information updating

In the previous subsections, we have shown that an introduction of information stickiness into Calvo pricing can fully explain the median persistence and volatility by searching for the *common* information delay, namely $\omega_j = \omega$ for all goods, j. In this subsection, we will briefly evaluate the obtained values of common information delay by comparing existing empirical macro studies on sticky information. Then, we relax our assumption of common information delay and consider *good-specific* information delays which account for the individual persistence or volatility. This consideration allows us to infer the differences of information delays across goods. We will then assess our results by comparing micro studies on price reviews in the U.S.

To evaluate common ω estimates, we first compare them with previous studies' estimates on information stickiness based on the aggregate inflation. Using the aggregate data on inflation over 1960:Q1 - 2007:Q2, DKT find that information delay, on average, is 7.1 months with 95 percent confidence intervals between 5.0 and 16.1 months. Knotek (2006) introduces information stickiness into the fixed menu cost model and finds the average duration between information updates to be 20.4 months over 1983:Q1 - 2005:Q4.[17] Thus, all of our common ω estimates (14 and 17 months

---

[17]Among many empirical studies on the pure sticky information model, Andrés, López-Salido, and Nelson (2005) estimate the average information duration to be 20 months and Kahn and Zhu (2006) find that the point estimates

from Bils and Klenow's data and 9.5 and 12 months from Nakamura and Steinsson's data) are in line with previous estimates based on aggregate inflation.

We now turn to a good specific ω by relaxing the assumption of common ω among goods. So far, our common values of ω was obtained to match the persistence and volatility for the median good. On the other hand, we can also obtain a *good specific* ω which matches the individual persistence or volatility for each good using the following procedure.

First, for each good $j(=1,2,...,165)$, we obtain good specific information delays from the persistence using

$$\min_{\omega_j \in [0,1)} [\hat{\alpha}_j - \alpha(\omega_j|\lambda_j, \rho)]^2,$$

where $\hat{\alpha}_j$ denotes the SAR estimate of the AR(4) model and $\alpha(\omega_j|\lambda_j, \rho)$ is the theoretical SAR given by $1 - (1 - \rho^{12})(1 - \lambda_j^{12})(1 - \omega_j^{12})(1 - (\omega_j\rho)^{12})$ evaluated at $\rho = 0.83$ from CKM and $\lambda_j = 1 - f_j$ from the frequency of price changes calculated by either Bils and Klenow (2004) or Nakamura and Steinsson (2007).

Second, we obtain good specific information delays from the volatility using

$$\min_{\omega_j \in [0,1)} [\hat{\sigma}_j - \sigma(\omega_j|\lambda_j, \rho, \beta)]^2,$$

where $\hat{\sigma}_j$ is the extracted standard deviation component of q_{it}^j due to time specific shocks normalized by the standard deviation of $\Delta \hat{S}_t$, while $\sigma(\omega_j|\lambda_j, \rho, \beta)$ is the predicted normalized standard deviation from the model under dual stickiness pricing.[18] We take $\rho = 0.83$, $\beta = 0.99$ and $\lambda_j = 1 - f_j$ from either Bils and Klenow (2004) or Nakamura and Steinsson (2007).

We now look at the distribution of good-specific average durations of information updates $1/(1 - \omega_j)$ based on the frequency of price changes from Bils and Klenow (2004). The upper and lower panels of Fig.9 show the relative histogram of information delays implied by the persistence and volatility, respectively.

The two kernel density estimates shown in the same figure, on the whole, suggest similarity between the two distributions. The median of the durations implied by persistence is 12.9 months while that of the durations implied by volatility is 16.6 months.[19] These values are close to the

---

of average duration range between 9 and 23 months.

[18]That is, $\sigma(\omega_j|\lambda_j, \rho, \beta)$ is the normalized standard deviation $\sigma_3(\omega_j, \lambda_j, \rho, \beta)$ evaluated at the fixed values of λ_j, ρ and β

[19]From the distribution implied by persistence after removing outliers, we obtain the standard deviation of 13.6, the skewness of 2.0, and the kurtosis of 5.7. On the other hand, we obtain the standard deviation of 15.6, the skewness of 1.5, and the kurtosis of 3.4 from the distribution implied by volatility.

average durations under the common ω assumption (14 months from persistence and 17 month from volatility).

With the frequency of price changes from Bils and Klenow (2004), we compute a fraction of goods in which persistence or volatility can be explained without information stickiness. The fraction is 11.5 percent from persistence matching while it is 6.1 percent from volatility matching. This computation implies that most goods need to have a positive good specific ω_j to fully explain good-level real exchange rate dynamics.

Next, we turn to Fig.10 which uses Nakamura and Steinsson's (2007) data on the frequency of price changes. Once again, the kernel density estimates suggest similarity of the two distributions. The median duration between information updates implied by persistence is 8.2 months while that implied by volatility is 11.9 months.[20] The fraction of goods that can match persistence or volatility without information stickiness has increased to 33.3 percent from persistence and 21.8 percent from volatility, due to the exclusion of sales. However, approximately more than two-thirds of goods still need to have a positive ω_j. Thus, the information stickiness remain important in explaining persistence and volatility with Nakamura and Steinsson's (2007) data.

Finally, we ask whether our results are, on the whole, consistent with evidence from micro studies on prices. Unfortunately, no micro studies provide directly comparable distribution of information delay among goods. However, survey results on price reviews done by firms may serve for our purpose. Fabiani, Druant, Hernando, Kwapil, Laudau, Loupias, Martins, Matha, Sabbatini, Stahl, and Stokman (2005) argue that the frequency of price reviews rather than price changes "could be related to the arrival of information." According to Fabiani et. al. (2005), when additional information on the state of the economy infrequently arrives, it is sensible for firms to review prices infrequently. In this sense, we can exploit survey results for price reviews.

Blinder, Canetti, Lebow, and Rudd (1998) surveyed U.S. firms about price setting behavior in the beginning of 1990s and their results for price reviews allow us to assess our distributions of average arrival of information. They ask a customary time interval (e.g.,daily, weekly, monthly, quarterly, and yearly) between price reviews for surveyed firms' most important product. Table 6 compares our distributions of durations of information updates with Blinder et. al. (1998) survey results. Overall, our distributions of duration between information updates seem to match the

---

[20]Descriptive statistics are as follows. From the distribution implied by persistence, we obtain the standard deviation of 10.8, the skewness of 2.3, and the kurtosis of 9.1. From the distribution implied by volatility, the corresponding statistics are 16.1, 1.9, and 4.7, respectively.

distribution of price reviews well. In particular, our results are close to their survey results when the frequency of price changes is taken from Nakamura and Steinsson (2007).

5 Conclusion

Using highly disaggregated price data from U.S. and Canadian cities, we have confirmed Kehoe and Midrigan's main finding that the standard Calvo-type sticky price model fails to explain the persistence and volatility of good-level real exchange rates. We found that this puzzling but stimulating result remains robust to a change from Bils and Klenow's data to Nakamura and Steinsson's data on the frequency of price changes. The robustness of their finding suggests that the model needs to be modified.

We offer a possible solution to this puzzle by extending the Kehoe-Midrigan such that only a fraction of firms have the up-to-date information when resetting prices. Due to the infrequent arrival of information, real exchange rates become more persistent and track the volatile nominal exchange rate even if price adjustment is relatively fast. Our model can explain both persistence and volatility within a reasonable range of average duration of information updates.

We have limited our attention to the implications of our model under many simplifying assumptions. Therefore, there are many promising avenues for future research. For example, what would happen to the prediction of our model if pricing complementarities are included? What would be the impact on good-level real exchange rate dynamics if the non-traded inputs in producing a good are included in the model?[21] We believe that answering these questions would help us further understand the dynamics of price adjustment within and across countries.

A The real exchange rate

From the intertemporal conditions (12) and (13), we obtain $\frac{P_{t+1}^* c_{t+1}^*}{P_{t+1} c_{t+1}} S_{t+1} = \frac{P_t^* c_t^*}{P_t c_t} S_t$ in each event in period $t+1$. Because q_t is defined as $\frac{S_t P_t^*}{P_t}$, it immediately follows that $q_{t+1} \frac{c_{t+1}^*}{c_{t+1}} = q_t \frac{c_t^*}{c_t} = q_{t-1} \frac{c_{t-1}^*}{c_{t-1}} = \cdots = q_0 \frac{c_0^*}{c_0} = \kappa$.

---

[21] See Crucini, Telmer, and Zachariadis (2005) for this line of research.

B The closed form solution to $\hat{x}_t(j,l)$

To derive the closed form solution to $\hat{x}_t(j,l)$, we use the closed form solution to $\hat{p}_{H,t}(j,l)$, given $\hat{\mu}_t$ follows AR(1). It has been already derived from (35) under an AR $\hat{\mu}_t$:

$$\hat{p}_{H,t}(j,l) = \left[\frac{\lambda_j \beta \rho}{1 - \lambda_j \beta \rho}\right] \hat{\mu}_t,$$

which implies

$$\mathbb{E}_{t-k-1}[\Delta \hat{p}_{H,t}(j,l)] = \left[\frac{\lambda_j \beta \rho}{1 - \lambda_j \beta \rho}\right] \mathbb{E}_{t-k-1}(\hat{\mu}_t - \hat{\mu}_{t-1})$$

$$= \left[\frac{\lambda_j \beta \rho}{1 - \lambda_j \beta \rho}\right] (\rho^{k+1} \hat{\mu}_{t-k-1} - \rho^k \hat{\mu}_{t-k-1}).$$

Using this result, we can express $\hat{x}_t(j,l)$ as $\hat{x}_{t-1}(j,l)$ and $\{\hat{\mu}_{t-k}\}_{k=0}^{\infty}$:

$$\hat{x}_t(j,l) = \omega_j \hat{x}_{t-1}(j,l) - \omega_j \mu_t + (1 - \omega_j) \left[\frac{\lambda_j \beta \rho}{1 - \lambda_j \beta \rho}\right] \hat{\mu}_t$$

$$+ \omega_j(1 - \omega_j) \sum_{k=0}^{\infty} \omega_j^k \left\{\left[\frac{\lambda_j \beta \rho}{1 - \lambda_j \beta \rho}\right] (\rho^{k+1} \hat{\mu}_{t-k-1} - \rho^k \hat{\mu}_{t-k-1}) + \rho^{k+1} \hat{\mu}_{t-k-1}\right\}.$$

Using a lag operator L, we can obtain

$$\hat{x}_t(j,l) = \omega_j \hat{x}_{t-1}(j,l) - \omega_j \hat{\mu}_t + (1 - \omega_j)\frac{\lambda_j \beta \rho}{1 - \lambda_j \beta \rho}\hat{\mu}_t$$

$$+ \omega_j(1 - \omega_j) \sum_{k=0}^{\infty} \omega_j^k \rho^k L^k \left\{\left[\frac{\lambda_j \beta \rho}{1 - \lambda_j \beta \rho}\right] [\rho - 1] + \rho\right\} \hat{\mu}_{t-1}.$$

Using $\sum_{k=0}^{\infty} \omega_j^k \rho^k L^k = (1 - \omega_j \rho L)^{-1}$ and arranging terms yields the closed form solution to $\hat{x}_t(j,l)$ given by (43).

C The proof of proposition 2

To prove Proposition 2, we use the first order difference equation for $\hat{p}_t(j,l)$ and $\hat{\mu}_t$ and an ARMA(1,∞) structure for $\hat{x}_t(j,l)$. We have

$$\hat{p}_t(j,l) = \lambda_j \hat{p}_{t-1}(j,l) - \lambda_j \hat{\mu}_t + (1 - \lambda_j)\hat{x}_t(j,l)$$

$$\hat{\mu}_t = \rho\hat{\mu}_{t-1} + \varepsilon_t$$

$$\hat{x}_t(j,l) = \omega_j \hat{x}_{t-1}(j,l) + a_j \hat{\mu}_t + \frac{b_j}{1 - \omega_j \rho L}\hat{\mu}_{t-1}$$

from (44), (31), and (43), respectively. We can rewrite the first and the third equations as follows:

$$\hat{p}_t(j,l) = -\frac{\lambda_j}{1 - \lambda_j L}\hat{\mu}_t + \frac{1 - \lambda_j}{1 - \lambda_j L}\hat{x}_t(j,l)$$

$$\hat{x}_t(j,l) = \frac{a_j}{1 - \omega_j L}\hat{\mu}_t + \frac{b_j}{(1 - \omega_j L)(1 - \omega_j \rho L)}\hat{\mu}_{t-1}.$$

We eliminate $\hat{x}_t(j,l)$ from these equations to get

$$(1 - \lambda_j L)(1 - \omega_j L)(1 - \omega_j \rho L)\hat{p}_t(j,l) = (1 - \lambda_j)a_j(1 - \omega_j \rho L)\hat{\mu}_t + (1 - \lambda_j)b_j\hat{\mu}_{t-1}$$
$$- \lambda_j(1 - \omega_j L)(1 - \omega_j \rho L)\hat{\mu}_t.$$

Arranging terms of the right hand side of the equation yields

$$(1 - \lambda_j L)(1 - \omega_j L)(1 - \omega_j \rho L)\hat{p}_t(j,l) = - \left[\lambda_j - (1 - \lambda_j)a_j\right]\hat{\mu}_t$$
$$+ \left[\lambda_j(\omega_j + \omega_j \rho) - (1 - \lambda_j)(\omega_j \rho a_j - b_j)\right]\hat{\mu}_{t-1}$$
$$- \lambda_j \omega_j^2 \rho \hat{\mu}_{t-2}.$$

Using the definition of $\theta_{j,0}$, $\theta_{j,1}$ and $\theta_{j,2}$ defined in Proposition 2, we get

$$(1 - \lambda_j L)(1 - \omega_j L)(1 - \omega_j \rho L)\hat{p}_t(j,l) = -\theta_{j,0}\hat{\mu}_t - \theta_{j,1}\hat{\mu}_{t-1} - \theta_{j,2}\hat{\mu}_{t-2}.$$

The left hand of the equation can be extended so that

$$(1 - \tilde{\phi}_{j,1}L - \tilde{\phi}_{j,2}L^2 - \tilde{\phi}_{j,3}L^3)\hat{p}_t(j,l) = -\theta_{j,0}\hat{\mu}_t - \theta_{j,1}\hat{\mu}_{t-1} - \theta_{j,2}\hat{\mu}_{t-2}.$$

Since the money growth rate follows an AR(1), $\hat{\mu}_t = (1 - \rho L)^{-1}\varepsilon_t$. Then,

$$(1 - \rho L)(1 - \tilde{\phi}_{j,1}L - \tilde{\phi}_{j,2}L^2 - \tilde{\phi}_{j,3}L^3)\hat{p}_t(j,l) = -\theta_{j,0}\varepsilon_t - \theta_{j,1}\varepsilon_{t-1} - \theta_{j,2}\varepsilon_{t-2}.$$

Arranging terms the left hand of the equation gives $\phi_{j,1}$, $\phi_{j,2}$, $\phi_{j,3}$, and $\phi_{j,4}$:

$$\hat{p}_t(j,l) = \sum_{r=1}^{4} \phi_{j,r}\hat{p}_t(j,l) - \sum_{r=0}^{2} \theta_{j,r}\varepsilon_{t-r}.$$

Analogously, we can derive the price index for good j of city l^*:

$$\hat{p}_t^*(j,l^*) = \sum_{r=1}^{4} \phi_{j,r}\hat{p}_t^*(j,l^*) - \sum_{r=0}^{2} \theta_{j,r}\varepsilon_{t-r}^*.$$

Because $\hat{q}_t(j,l,l^*) = \hat{p}_t^*(j,l^*) - \hat{p}_t(j,l)$, we obtain (45).

Finally, note that the coefficient of $\hat{p}_t(j,l)$ is

$$(1 - \rho L)(1 - \tilde{\phi}_{j,1}L - \tilde{\phi}_{j,2}L^2 - \tilde{\phi}_{j,3}L^3) = (1 - \rho L)(1 - \lambda_j L)(1 - \omega_j L)(1 - \omega_j \rho L).$$

It implies that the SAR $\sum_{r=1}^{4} \phi_{j,r}$ is equal to $1 - (1 - \rho)(1 - \lambda_j)(1 - \omega_j)(1 - \omega_j \rho)$. Because the AR coefficients are the same between the same type of good j, it proves Proposition 2.

D The Detailed Derivation of Transformation from Monthly to Annual Specification

This appendix shows how we transform a monthly specification into the one which has non-zero AR coefficients for multiples of 12 month lags and finite MA terms with the remaining AR coefficients equal to zero. Table 1 summarizes the obtained results before and after transformations. The transformations lead us to estimate the model via the annual data.

We have already shown the transformation results of Calvo pricing with $\rho = 0$ in the main text. In what follows, we will show the derivation of Calvo pricing with $\rho > 0$ and dual stickiness pricing.

Calvo pricing ($\rho > 0$) First, we can rewrite the first order difference equation (40) as

$$\hat{q}_t(j,l,l^*) = \lambda_j^{12}\hat{q}_{t-1}(j,l,l^*) + \theta_j\Lambda_j(L)\hat{S}_t = \frac{\theta_j\Lambda_j(L)}{1 - \lambda_j^{12}L^{12}}\Delta\hat{S}_t, \tag{A1}$$

Second, since $\Delta\hat{S}_t = \hat{\mu}_t - \hat{\mu}_t^*$, it immediately follows that

$$\Delta\hat{S}_t = \rho\Delta\hat{S}_{t-1} + \eta_t = \frac{R(L)}{1 - \rho^{12}L^{12}}\eta_t, \tag{A2}$$

where $R(L) = \sum_{r=0}^{11}\rho^r L^r$. Substituting (A2) into (A1) yields:

$$\hat{q}_t(j,l,l^*) = (\lambda_j^{12} + \rho^{12})\hat{q}_{t-12}(j,l,l^*) - \lambda_j^{12}\rho^{12}\hat{q}_{t-24}(j,l,l^*) + \theta_j\Lambda_j(L)R(L)\eta_t, \tag{A3}$$

which produces an ARMA(24,22).[22] The AR parameters are non-zero only if the lags are multiples of 12. Moreover, the length of the MA terms is now finite and of order 22 in this specific ARMA process. Intuitively, this transformation is made possible because $\hat{q}_t(j,l,l^*)$ is the first order difference equation and the driving force $\Delta\hat{S}_t$ follows an AR(1) process. Remarkably, this monthly ARMA(24,22) becomes ARMA(2,1) in terms of annually sampled data.

Dual stickiness pricing A similar transformation is possible in dual stickiness pricing. The next proposition summarizes the transformation result.

Proposition A1. *In dual stickiness pricing with ρ, λ_j, and $\omega_j \in (0,1)$, the ARMA(4,2) process characterized by (45) has an equivalent expression of the following ARMA(48,46) process:*

$$\hat{q}_t(j,l,l^*) = \sum_{r=1}^{4}\Phi_{j,r}\hat{q}_{t-12r}(j,l,l^*) + \Theta_j(L)\eta_t, \tag{A4}$$

---

[22]It is because both $\Lambda_j(L)$ and $R(L)$ have the power of L of 11 in maximum.

where

$$\Phi_{j,1} = \tilde{\Phi}_{j,1} + \rho^{12}, \qquad \tilde{\Phi}_{j,1} = \lambda_j^{12} + \omega_j^{12} + (\omega_j\rho)^{12}$$

$$\Phi_{j,2} = \tilde{\Phi}_{j,2} - \tilde{\Phi}_{j,1}\rho^{12}, \qquad \tilde{\Phi}_{j,2} = -[\lambda_j^{12}\omega_j^{12} + (\lambda_j^{12} + \omega_j^{12})\omega_j^{12}\rho^{12}]$$

$$\Phi_{j,3} = \tilde{\Phi}_{j,3} - \tilde{\Phi}_{j,2}\rho^{12}, \qquad \tilde{\Phi}_{j,3} = \lambda_j^{12}\omega_j^{24}\rho^{12}$$

$$\Phi_{j,4} = -\tilde{\Phi}_{j,3}\rho^{12}$$

$$\Theta_j(L) = \left\{ (1 - \omega_j^{12}L^{12})(1 - (\omega_j^{12}\rho)^{12}L^{12})\lambda_j\Lambda_j(L)R(L) \right.$$

$$\left. -(1 - \lambda_j)\Lambda_j(L)\Omega_j(L)R(L)\left[(1 - (\omega_j\rho)^{12}L^{12})a_j + b_jL(1 + \Omega_j^R(L))\right] \right\}$$

$$\Omega_j(L) = \sum_{r=0}^{11}\omega_j^r L^r, \quad \Omega_j^R(L) = \sum_{r=1}^{11}(\omega_j\rho)^r L^r.$$

Proof. To consider the transformation under dual stickiness pricing, note that (43) can be rewritten as

$$\hat{x}_t(j,l) = \omega_j\hat{x}_{t-1}(j,l) + a_j\hat{\mu}_t + \frac{b_jL}{1 - \omega_j\rho L}\hat{\mu}_t,$$

using a lag operator L. This equation has an infinite MA term because the third term of the right hand side has $(1 - \omega_j\rho L)^{-1}\hat{\mu}_t$. We first work on this term.

The infinite MA form $(1 - \omega_j\rho L)^{-1}\hat{\mu}_t$ is

$$(1 - \omega_j\rho L)^{-1}\hat{\mu}_t \quad = \hat{\mu}_t \quad + \quad \sum_{r=1}^{11}(\omega_j\rho)^r\hat{\mu}_{t-r}$$

$$+(\omega_j\rho)^{12}\hat{\mu}_{t-12} \quad + \quad (\omega_j\rho)^{12}\sum_{r=1}^{11}(\omega_j\rho)^r\hat{\mu}_{t-r-12}$$

$$+(\omega_j\rho)^{24}\hat{\mu}_{t-24} \quad + \quad (\omega_j\rho)^{24}\sum_{r=1}^{11}(\omega_j\rho)^r\hat{\mu}_{t-r-24} + \cdots.$$

Collecting terms by columns yields

$$(1 - \omega_j\rho L)^{-1}\hat{\mu}_t = (1 + (\omega_j\rho)^{12}L^{12} + (\omega_j\rho)^{24}L^{24} + \cdots)\hat{\mu}_t$$

$$+(1 + (\omega_j\rho)^{12}L^{12} + (\omega_j\rho)^{24}L^{24} + \cdots)\sum_{r=1}^{11}(\omega_j\rho)^r\hat{\mu}_{t-r}$$

$$= \frac{1}{1 - (\omega_j\rho)^{12}L^{12}}\hat{\mu}_t + \frac{1}{1 - (\omega_j\rho)^{12}L^{12}}\sum_{r=1}^{11}(\omega_j\rho)^r L^r\hat{\mu}_t$$

$$= \frac{1 + \Omega_j^R(L)}{1 - (\omega_j\rho)^{12}L^{12}}\hat{\mu}_t,$$

where $\Omega_j^R(L) = \sum_{r=1}^{11}(\omega_j\rho)^r L^r$.

Using this result, we obtain the first order difference equation for $\hat{x}_t(j, l)$:

$$\hat{x}_t(j, l) = \omega_j \hat{x}_{t-1}(j, l) + \left[a_j + \frac{b_j L(1 + \Omega_j^R(L))}{1 - (\omega_j \rho)^{12} L^{12}}\right] \hat{\mu}_t.$$

Equivalently, by repeated substitutions,

$$\hat{x}_t(j, l) = \omega_j^{12} \hat{x}_{t-12}(j, l) + \left[a_j + \frac{b_j L(1 + \Omega_j^R(L))}{1 - (\omega_j \rho)^{12} L^{12}}\right] \Omega_j(L) \hat{\mu}_t, \tag{A5}$$

where $\Omega_j(L) = \sum_{r=0}^{11} \omega_j^r L^r$.

Similarly, the equation for the good j price index is the first order difference equation given by (44). It implies

$$\hat{p}_t(j, l) = \lambda_j^{12} \hat{p}_{t-12}(j, l) - \lambda_j \Lambda_j(L) \hat{\mu}_t + (1 - \lambda_j) \Lambda_j(L) \hat{x}_t(j, l). \tag{A6}$$

Substituting (A5) into (A6) yields

$$\begin{aligned}
\hat{p}_t(j, l) = &-\frac{\lambda_j \Lambda_j(L)}{1 - \lambda_j^{12} L^{12}} \hat{\mu}_t \\
&+ \frac{(1 - \lambda_j) \Lambda_j(L) \Omega_j(L) \left[(1 - (\omega_j \rho)^{12} L^{12}) a_j + b_j L(1 + \Omega_j^R(L))\right]}{(1 - \lambda_j^{12} L^{12})(1 - \omega_j^{12} L^{12})(1 - (\omega_j \rho)^{12} L^{12})} \hat{\mu}_t.
\end{aligned} \tag{A7}$$

Analogously, we can obtain a similar equation for $\hat{p}_t^*(j, l^*)$. Then, noting that $\hat{q}_t(j, l, l^*) = \hat{p}_t^*(j, l^*) - \hat{p}_t(j, l)$ and $\Delta \hat{S}_t = \hat{\mu}_t - \hat{\mu}_t^*$, we can obtain the following equation for the good-level real exchange rate:

$$\begin{aligned}
\hat{q}_t(j, l, l^*) = &\frac{\lambda_j \Lambda_j(L)}{1 - \lambda_j^{12} L^{12}} \Delta \hat{S}_t \\
&- \frac{(1 - \lambda_j) \Lambda_j(L) \Omega_j(L) \left[(1 - (\omega_j \rho)^{12} L^{12}) a_j + b_j L(1 + \Omega_j^R(L))\right]}{(1 - \lambda_j^{12} L^{12})(1 - \omega_j^{12} L^{12})(1 - (\omega_j \rho)^{12} L^{12})} \Delta \hat{S}_t
\end{aligned}$$

Arranging the terms and using $\Delta \hat{S}_t = (1 - \rho^{12} L^{12})^{-1} R(L) \eta_t$, we obtain

$$\begin{aligned}
&(1 - \lambda_j^{12} L^{12})(1 - \omega_j^{12} L^{12})(1 - (\omega_j \rho)^{12} L^{12}) \hat{q}_t(j, l, l^*) \\
&= \left\{ (1 - \omega^{12} L^{12})(1 - (\omega_j \rho)^{12} L^{12}) \lambda_j \Lambda_j(L) R(L) \right. \\
&\left. - (1 - \lambda_j) \Lambda_j(L) \Omega_j(L) R(L) \left[(1 - (\omega_j \rho)^{12} L^{12}) a_j + b_j L(1 + \Omega_j^R(L))\right] \right\} \frac{\eta_t}{1 - \rho^{12} L^{12}}.
\end{aligned}$$

The terms inside the curly bracket gives $\Theta_j(L)$. Moreover, the first line of the terms has non-zero coefficient for L^{46}, because $(1 - \omega^{12} L^{12})(1 - (\omega_j \rho)^{12} L^{12})$ have a non-zero coefficient for L^{24} and

$\Lambda_j(L)R(L)$ have a non-zero coefficient for L^{22}. Since the second line of the terms inside the curly brackets have L^{45}, the maximum power for L is 46.

By multiplying both sides of the equation by $(1 - \rho^{12}L^{12})$ we obtain from (A2) to get

$$(1 - \rho^{12}L^{12})(1 - \tilde{\Phi}_{j,1}L^{12} - \tilde{\Phi}_{j,2}L^{24} - \tilde{\Phi}_{j,3}L^{36})\hat{q}_t(j,l,l^*) = \Theta_j(L)\eta_t,$$

which gives us $\Phi_{j,1}$, $\Phi_{j,2}$, $\Phi_{j,3}$ and $\Phi_{j,4}$. $\qquad\square$

The implications of Proposition A1 are as follows. First, the number of AR parameters are limited to four and these four parameters are the coefficients on lags of 12, 24, 36, and 48 months. Thus, the autoregressive part of the model has a form of autoregression on the past values of the real exchange rates at annual frequencies. Second, if the AR part has the restriction described above and if the maximum order of MA coefficients is 46, dual stickiness pricing with ρ, λ_j, and $\omega_j \in (0,1)$ can be written only with this representation. Third, this ARMA(48,46) process becomes ARMA(4,3) in terms of annually sampled data. Finally, under the representation, we can also show that the SAR is given by

$$\alpha_j = \sum_{r=1}^{4} \Phi_{j,r} = 1 - (1 - \rho^{12})(1 - \lambda_j^{12})(1 - \omega_j^{12})(1 - (\omega_j\rho)^{12}), \tag{A8}$$

which is increasing in λ_j, ω_j and ρ.

E   The long-run value of a good-level real exchange rate

This appendix shows the long-run value of $q_t(j,l,l^*)$. In what follows, we use variables without time subscript to denote the steady state value.

Consider the steady state value of the price of good j in city l. In the steady state, U.S. firms set prices such that

$$P_H(j,l) = \frac{\theta}{\theta - 1}W = \chi\frac{\theta}{\theta - 1}M.$$

Here, we used (16). Canadian firms choose prices such that

$$P_F(j,l) = \frac{\theta}{\theta - 1}(1 + \tau(j,l))SW^* = \chi\frac{\theta}{\theta - 1}\kappa(1 + \tau(j,l))M.$$

because of (17) and (19). Therefore, the price of good j in city l is

$$P(j,l) = \frac{\chi}{2}\frac{\theta}{\theta - 1}[1 + \kappa^{1-\theta}(1 + \tau(j,l))^{1-\theta}]^{\frac{1}{1-\theta}}M. \tag{A9}$$

By similar argument, we can derive $P^*(j, l^*)$ as follows:

$$P^*(j, l^*) = \frac{\chi}{2} \frac{\theta}{\theta - 1} [\kappa^{-(1-\theta)} + (1 + \tau(j, l))^{1-\theta}]^{\frac{1}{1-\theta}} M^*. \tag{A10}$$

Given the good-level real exchange rate for good j of a city pair between l and l^* is given by $q(j, l, l^*) = SP(j, l)/P(j, l^*)$, the equations (19), (A9), and (A10) imply

$$q(j, l, l^*) = \frac{[1 + \kappa^{1-\theta}(1 + \tau(j, l^*))^{1-\theta}]^{\frac{1}{1-\theta}}}{[1 + \kappa^{1-\theta}(1 + \tau(j, l))^{1-\theta}]^{\frac{1}{1-\theta}}}. \tag{A11}$$

Thus, the long-run value of a good-level of real exchange rate depends on the city pair.

References

AHLIN, C., AND M. SHINTANI (2007): "Menu Costs and Markov Inflation: A Theoretical Revision with New Evidence," *Journal of Monetary Economics*, 54(3), 753–784.

ANDRÉS, J., J. D. LÓPEZ-SALIDO, AND E. NELSON (2005): "Sticky Price Models and the Natural Rate of Hypothesis," *Journal of Monetary Economics*, 52(5), 1025–1053.

ANDREWS, D. W. K., AND H.-Y. CHEN (1994): "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics*, 12(2), 187–204.

ARELLANO, M., AND S. BOND (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economics Studies*, 58(2), 277–297.

BILS, M., AND P. J. KLENOW (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112(5), 947–985.

BLINDER, A. S., E. R. D. CANETTI, D. E. LEBOW, AND J. B. RUDD (1998): *Asking about Prices: A New Approach to Understanding Price Stickiness*. Russell Sage Foundation, New York.

CALVO, G. (1983): "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12(3), 383–398.

CHARI, V. V., P. J. KEHOE, AND E. R. MCGRATTAN (2002): "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?," *Review of Economics Studies*, 69(3), 533–563.

CLARK, T. E. (2006): "Disaggregate Evidence on the Persistence of Consumer Price Inflation," *Journal of Applied Econometrics*, 21(5), 563–587.

CRUCINI, M. J., AND M. SHINTANI (2008): "Persistence in Law of One Price Deviations: Evidence from Micro-data," *Journal of Monetary Economics,* forthcoming.

CRUCINI, M. J., AND C. TELMER (2007): "Microeconomic Sources of Real Exchange Rate Variability," unpublished manuscript.

CRUCINI, M. J., C. I. TELMER, AND M. ZACHARIADIS (2005): "Understanding European Real Exchange Rates," *American Economic Review*, 95(3), 724–738.

DUPOR, B., T. KITAMURA, AND T. TSURUGA (2008): "Integrating Sticky Prices and Sticky Information," unpublished manuscript.

ENGEL, C., AND J. H. ROGERS (2004): "European Market Integration after Euro," *Economic Policy*, 19, 347–384.

FABIANI, S., M. DRUANT, I. HERNANDO, C. KWAPIL, B. LAUDAU, C. LOUPIAS, F. MARTINS, T. Y. MATHA, R. SABBATINI, H. STAHL, AND A. C. J. STOKMAN (2005): "The Pricing Behaviour of Firms in the Euro Area New Survey Evidence," ECB working paper No 535.

KAHN, H., AND Z. ZHU (2006): "Estimates of the Sticky-Information Phillips Curve for the United States," *Journal of Money, Credit and Banking*, 38(1), 195–208.

KEHOE, P. J., AND V. MIDRIGAN (2007): "Sticky Prices and Sectoral Real Exchange Rates," Federal Reserve Bank of Minneapolis Working Paper No 656.

KNOTEK, E. S. (2006): "A Tale of Two Rigidities: Sticky Prices in a Sticky Information Environment," Federal Reserve Bank of Kansas City Research Working Paper 06-15.

MANKIW, N. G., AND R. REIS (2002): "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117(4), 1295–1328.

——— (2006): "Pervasive Stickiness," *American Economic Review*, 96(2), 164–169.

NAKAMURA, E., AND J. STEINSSON (2007): "Five Facts About Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics* forthcoming.

PARSLEY, D. C., AND S.-J. WEI (2007): "A Prism into the PPP Puzzles: The Micro-foundations of Big Mac Real Exchange Rates," *Economic Journal*, 117, 1336–1356.

ROGERS, J. H. (2007): "Monetary Union, Price Level Convergence, and Inflation: How close is Europe to the USA?," *Journal of Monetary Economics*, 54(3), 785–796.

SIMS, C. A. (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3), 665–690.

WOODFORD, M. (2003): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, chap. 1. Princeton University Press.

YUN, T. (1996): "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles," *Journal of Monetary Economics*, 37(2), 345–370.

Figure 1: Persistence without information delay: function of money growth parameter(ρ) and Calvo parameter (λ_j)
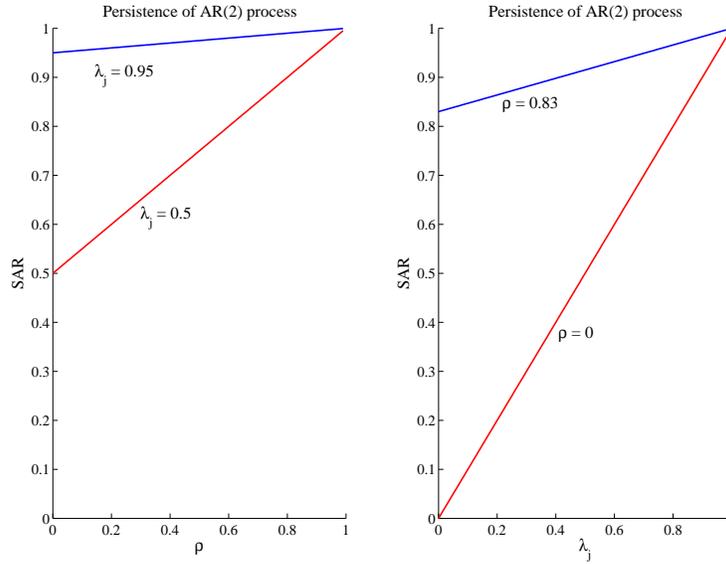

Figure 2: Volatility without information delay: function of money growth parameter(ρ) and Calvo parameter(λ_j)
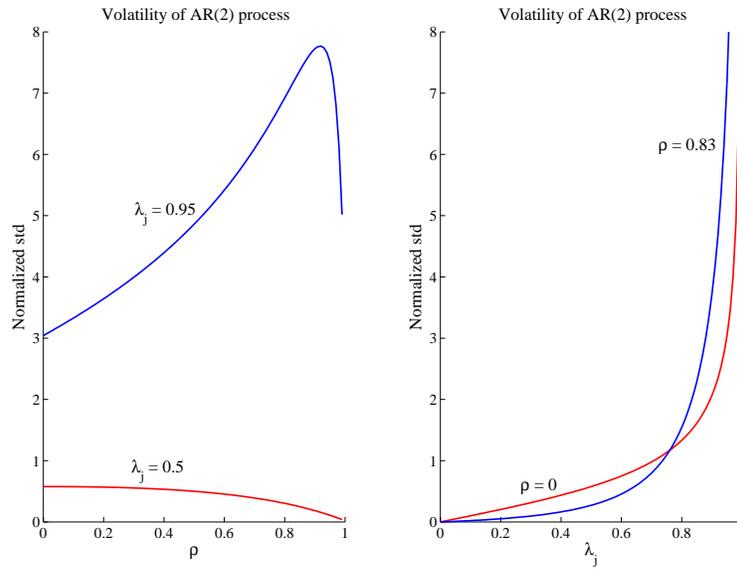
NOTES: The discount factor β is 0.99.

Figure 3: Persistence with information delay: function of information stickiness parameter(ω_j) and Calvo parameter(λ_j)
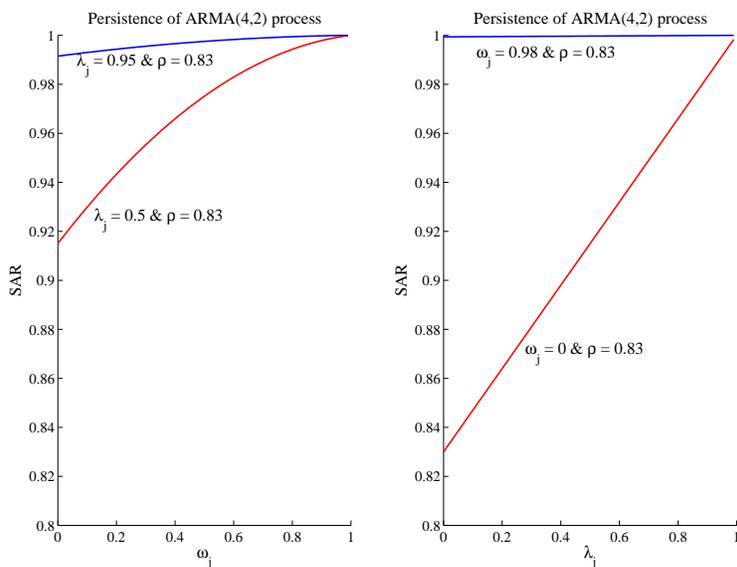
Figure 4: Volatility with information delay: function of information stickiness parameter(ω_j) and Calvo parameter(λ_j)
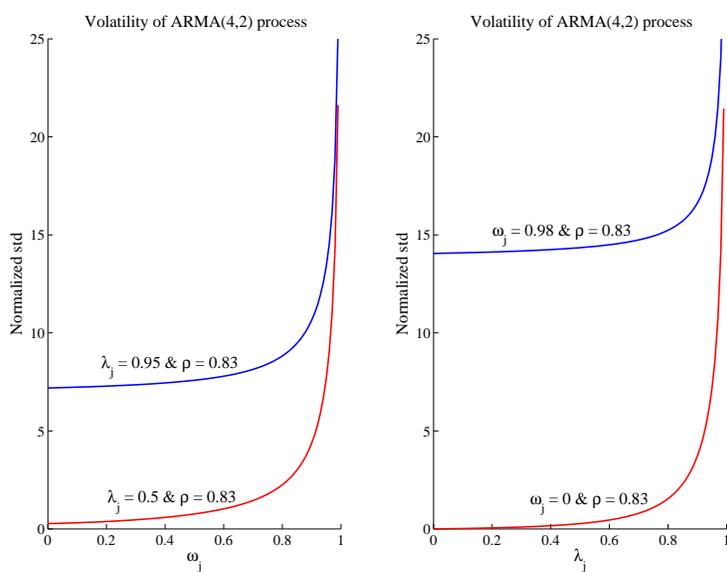
NOTES: The discount factor β is 0.99.

Figure 5: The cross-border city pairs in the U.S. and Canada

Figure 6: Empirical distribution of LOP deviations in 1990 and 2005
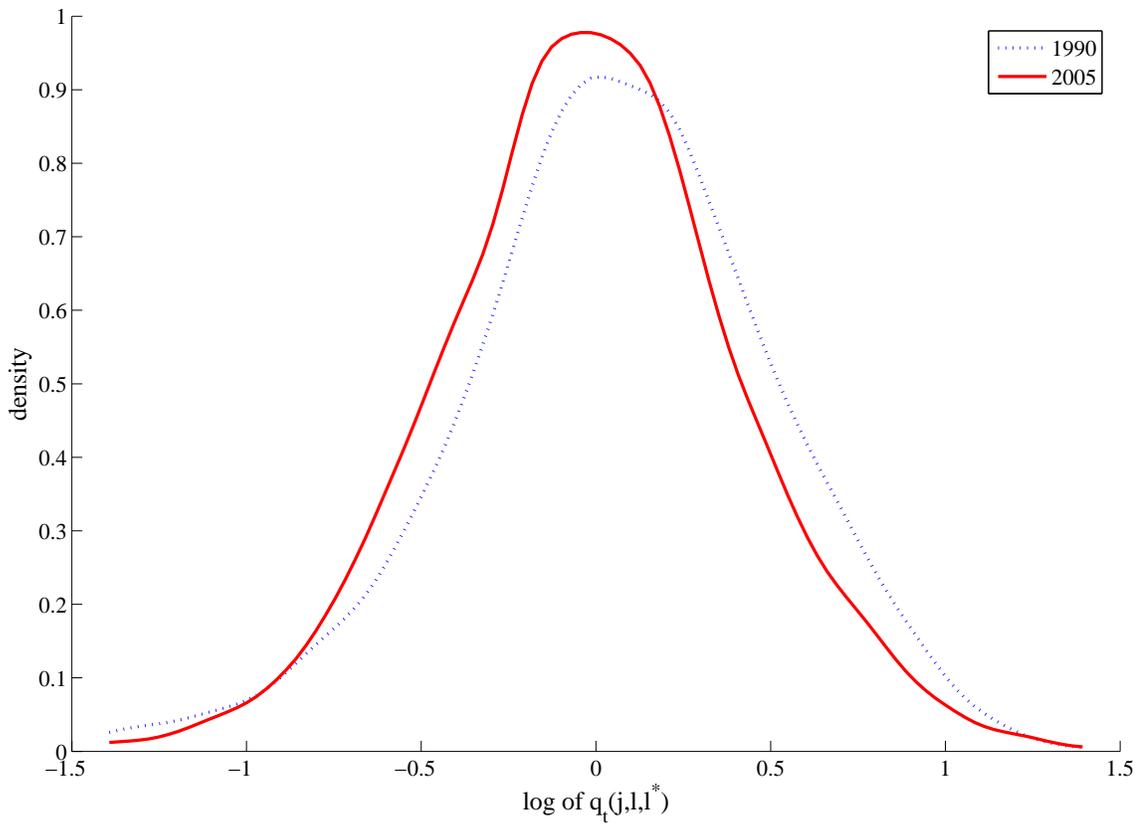
NOTES: Kernel density estimates.

Figure 7: Real exchange rate persistence and price stickiness: Bils and Klenow (2004)
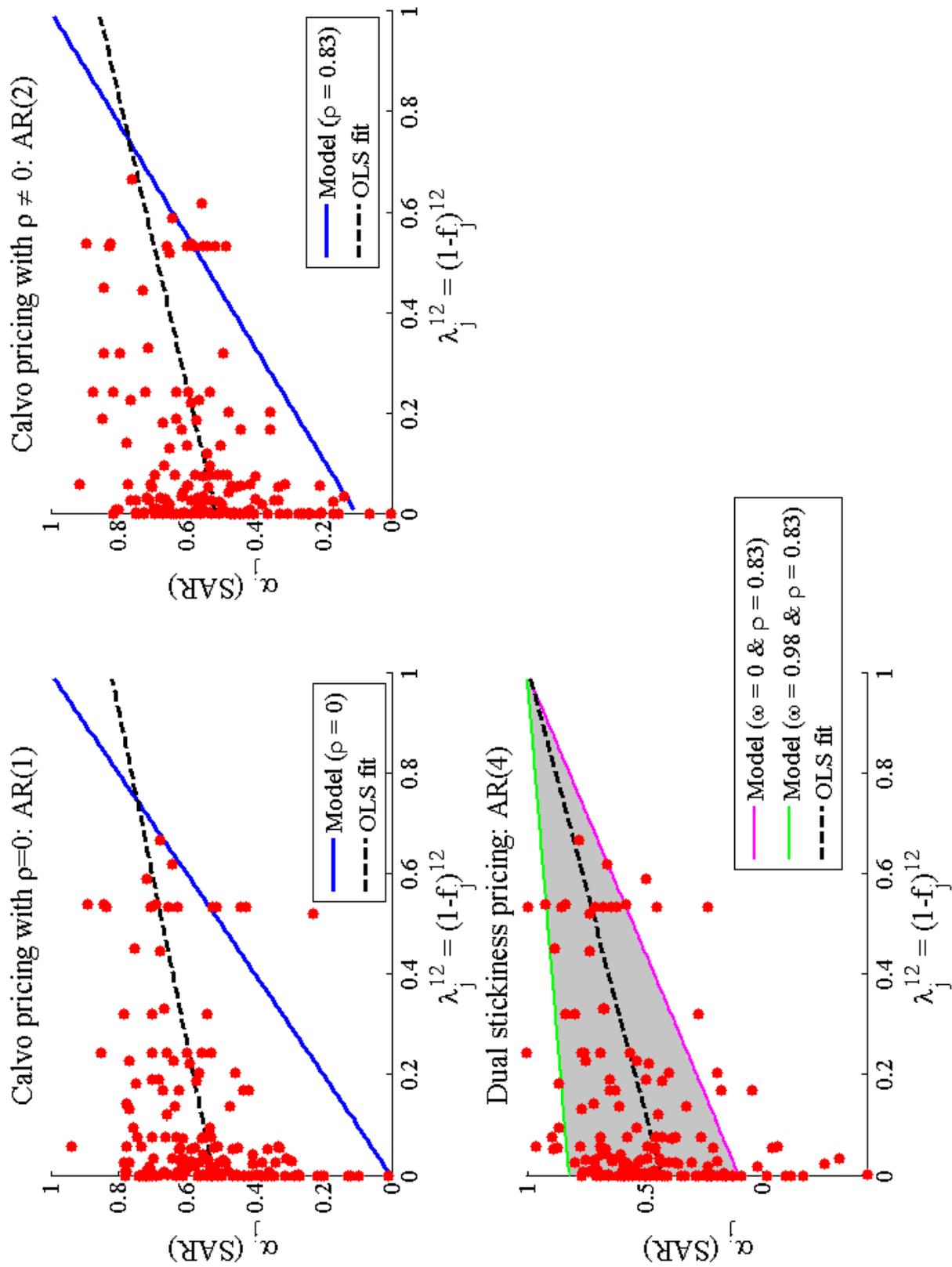
Figure 8: Real exchange rate persistence and price stickiness: Nakamura and Steinsson (2007)
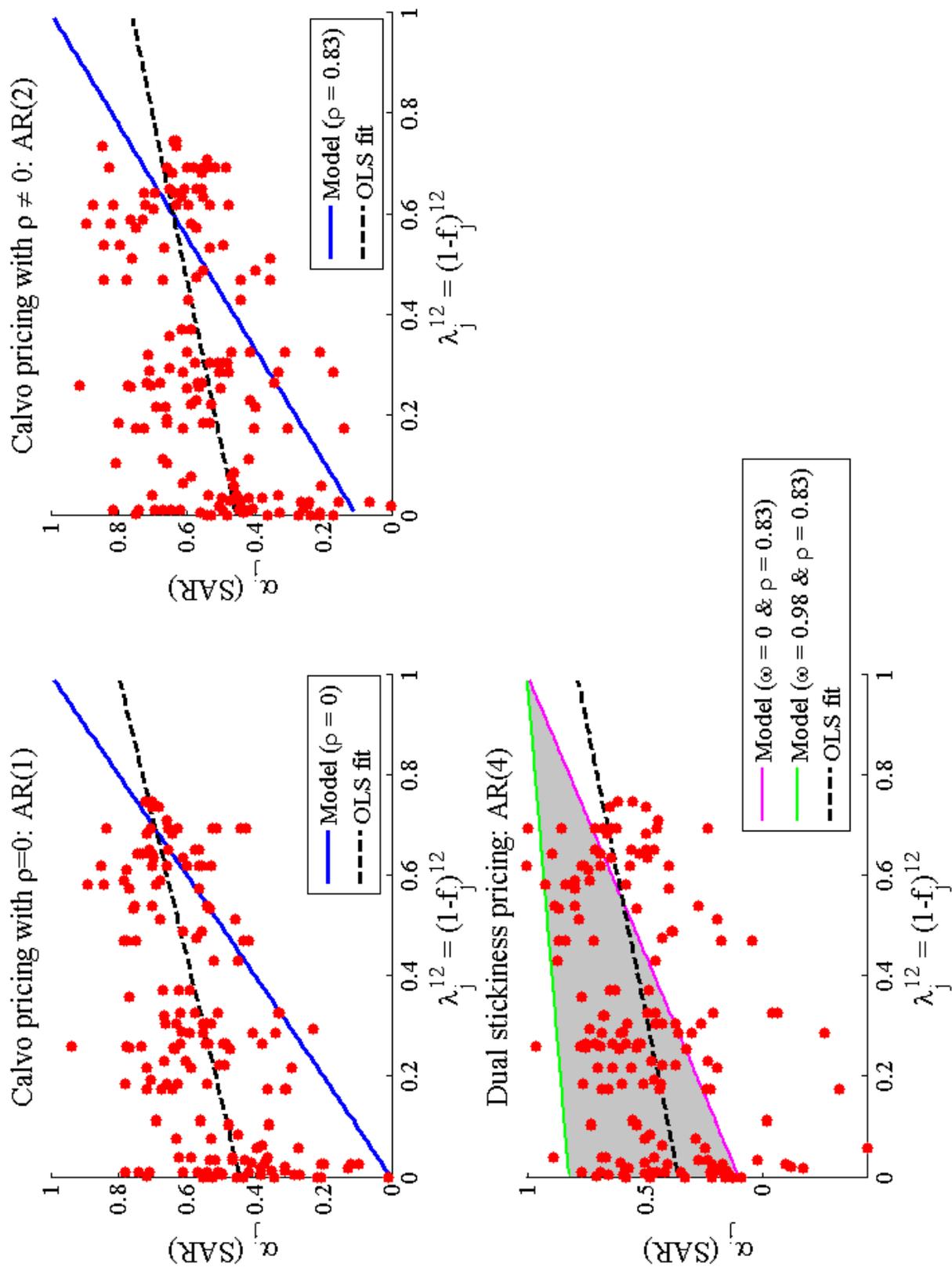
Figure 9: Empirical distribution of good-specific information delays $1/(1 - \omega_j)$: Bils and Klenow (2004)
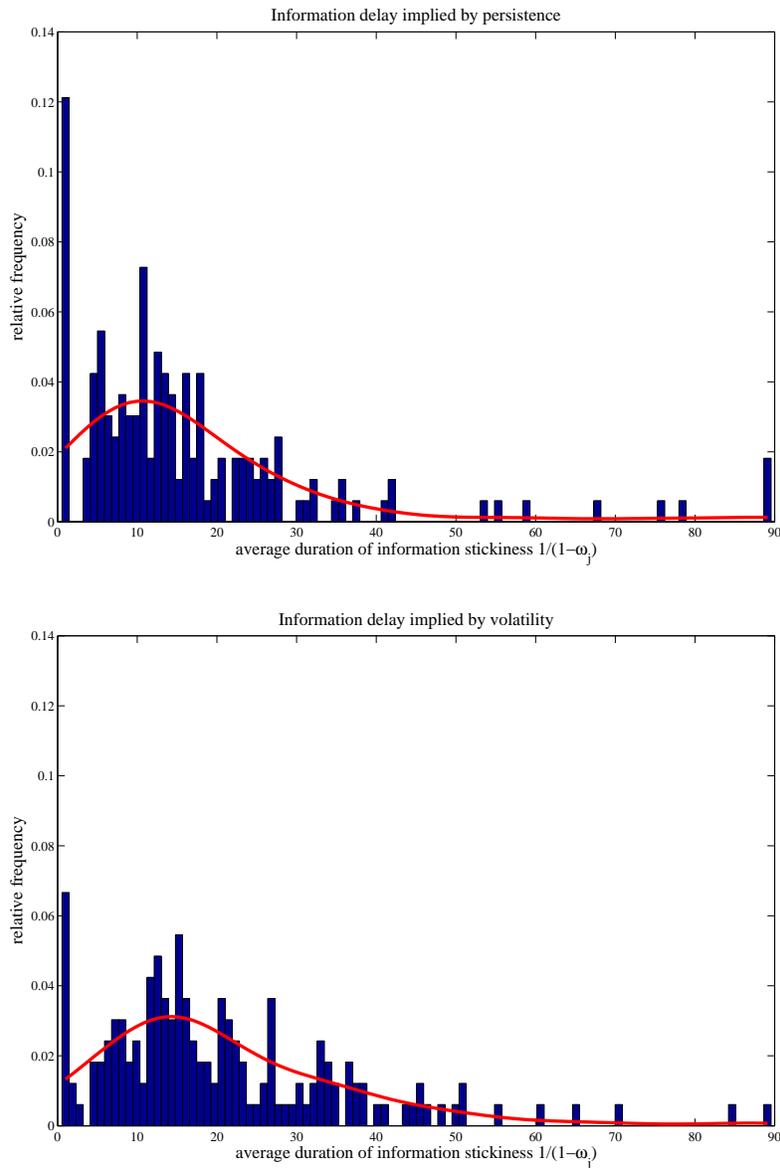
NOTES: The upper panel shows the relative histogram of average information delay $1/(1 - \omega_j)$ where $\omega_j \in [0, 1)$ for each good j is obtained by minimizing t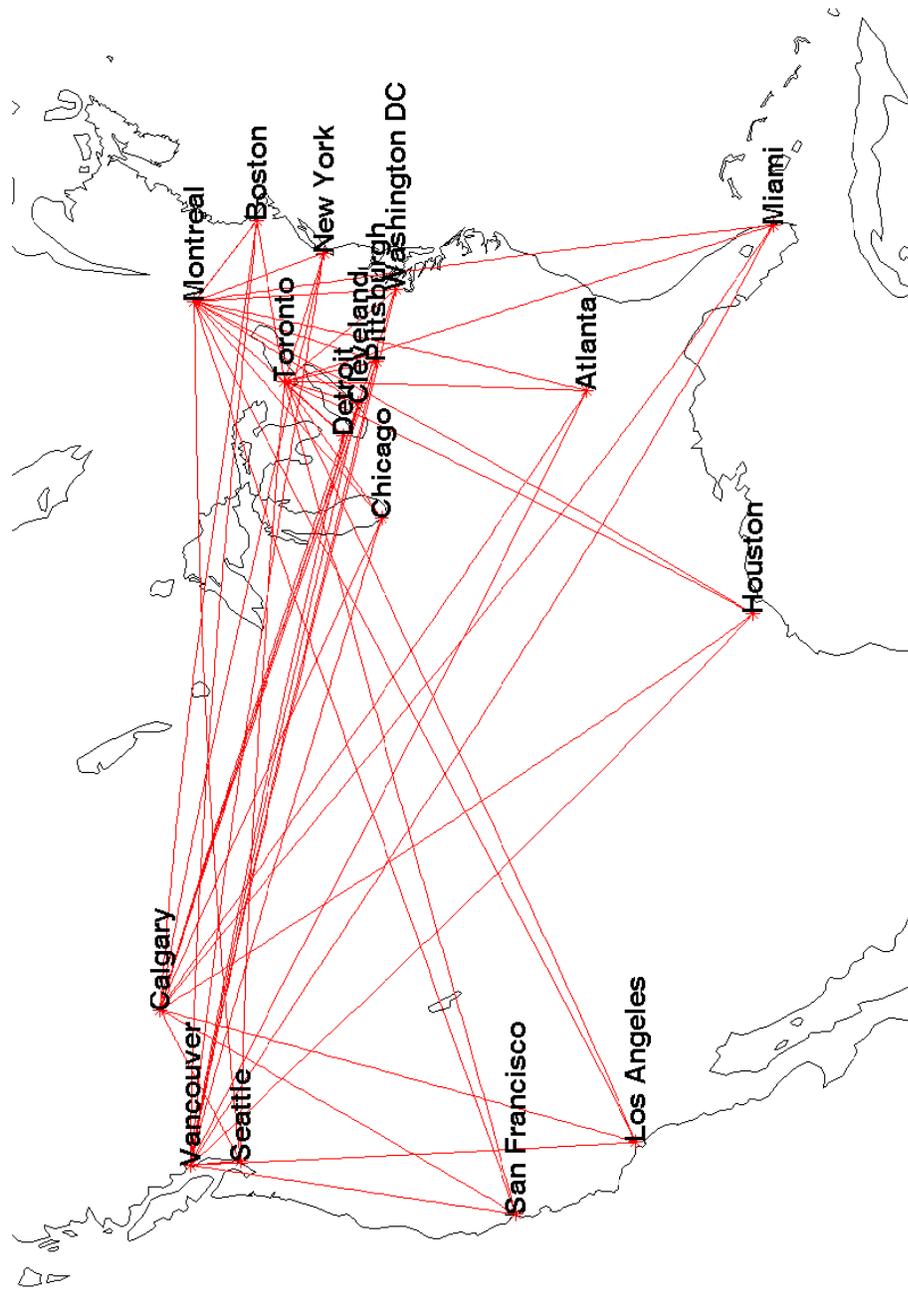he distance between the observed SAR and theoretical prediction from dual stickiness pricing using $\lambda_j = 1 - f_j$ from Bils and Klenow (2004) and $\rho = 0.83$. The smoothed lines are kernel density estimates. The lower panel shows the distribution when each ω_j is obtained by minimizing the distance between the observed volatility and theoretical prediction using $\lambda_j = 1 - f_j$, $\rho = 0.83$ and $\beta = 0.99$.

Figure 10: Empirical distribution of good-specific information delays $1/(1 - \omega_j)$: Nakamura and Steinsson (2007)

Information delay implied by persistence

Information delay implied by volatility
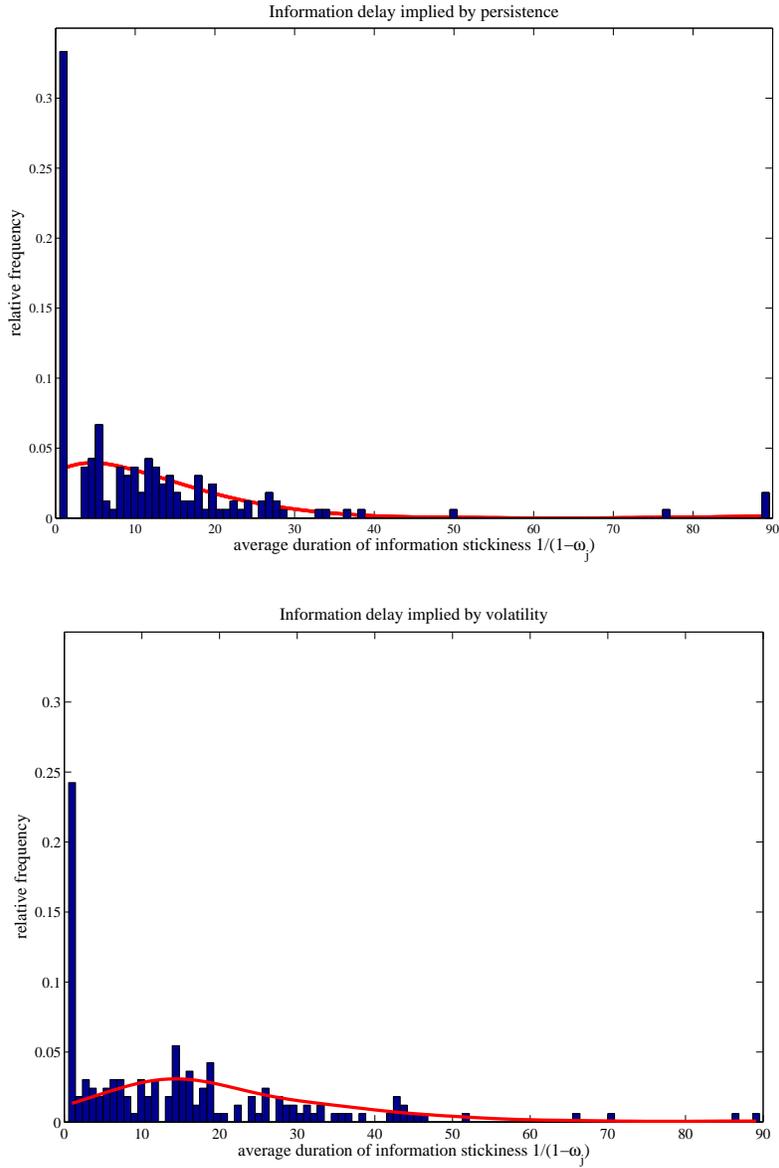
NOTES: See the notes of Fig.9. Nakamura and Steinsson's (2007) frequency of price changes, in stead of Bils and Klenow's (2004), is used for $\lambda_j = 1 - f_j$ in the computation of the theoretical value.

Table 1: Summary of transformations from monthly to annual specification

	Monthly specification	Annual specification
Calvo (ρ=0)	$\hat{q}_t(j,l,l^*) = \lambda_j \hat{q}_{t-1}(j,l,l^*) - \lambda_j \eta_t$	$\hat{q}_t(j,l,l^*) = \lambda_j^{12} \hat{q}_{t-12}(j,l,l^*) - \lambda_j \Lambda_j(L)\eta_t$
Calvo ($\rho > 0$)	$\hat{q}_t(j,l,l^*) = (\lambda_j + \rho)\hat{q}_{t-1}(j,l,l^*)$ $-\lambda_j \rho \hat{q}_{t-2}(j,l,l^*) - \theta_j \eta_t$	$\hat{q}_t(j,l,l^*) = (\lambda_j^{12} + \rho^{12})\hat{q}_{t-12}(j,l,l^*)$ $-\lambda_j^{12}\rho^{12}\hat{q}_{t-24}(j,l,l^*) - \lambda_j \Lambda_j(L)R(L)\eta_t$
Dual stickiness	$\hat{q}_t(j,l,l^*) = \sum_{r=1}^{4} \phi_{j,r}\hat{q}_{t-r}(j,l,l^*)$ $+ \sum_{r=0}^{2} \theta_{j,r}\eta_{t-r}$	$\hat{q}_t(j,l,l^*) = \sum_{r=1}^{4} \Phi_{j,r}\hat{q}_{t-12r}(j,l,l^*)$ $+\Theta_j(L)\eta_t$

NOTES: The left panel shows the original monthly ARMA processes which are in the main text. The right panel shows corresponding conversions such that autoregressive coefficients are non-zero only if the lags are multiples of 12 and that moving average terms are finite. These conversions allow us to estimate the original monthly ARMA process with annually sampled data. The autoregressive parameters $\Phi_{j,r}$ and moving average polynomials, $\Lambda_j(L)$, $R(L)$ and $\Theta_j(L)$ are given in Appendix D.

Table 2: Proportions of explained persistence of good-level real exchange rates: Bils and Klenow (2004)

Calvo pricing with various ρ								
ρ	0	0.3	0.5	0.7	0.9	0.95	0.98	0.946
Theory/Data	0.059	0.058	0.058	0.088	0.634	1.043	1.425	1.000

Dual stickiness pricing with $\rho = 0.83$								
ω	0	0.3	0.5	0.7	0.9	0.95	0.98	0.930
Theory/Data	0.306	0.323	0.323	0.350	0.792	1.209	1.529	1.000

NOTES: Numbers are median ratios of the theoretical persistence, predicted by Bils and Klenow (2004), to observed persistence measured by SAR estimated from real exchange rate data. Theoretical persistence for the upper panel is the SAR for various ρ when Calvo pricing is used. Theoretical persistence for the lower panel is the SAR for various common ω with $\rho = 0.83$ when dual stickiness pricing is used. Median SAR estimates for AR(1), AR(2) and AR(4) models are 0.563, 0.568, and 0.508, respectively. The last column of each panel shows the value of ρ and ω, giving the median ratio closest to one.

Table 3: Proportions of explained persistence of good-level real exchange rates: Nakamura and Steinsson (2007)

Calvo pricing with various ρ

ρ	0	0.3	0.5	0.7	0.9	0.95	0.98	0.923
Theory/Data	0.484	0.505	0.506	0.549	0.922	1.226	1.522	1.000

Dual stickiness pricing with $\rho = 0.83$

ω	0	0.3	0.5	0.7	0.9	0.95	0.98	0.895
Theory/Data	0.664	0.659	0.660	0.681	1.015	1.329	1.619	1.000

NOTES: See the notes of Table 2. Nakamura and Steinsson's (2007) data, instead of Bils and Klenow's (2004), is used for the computation of the theoretical value.

Table 4: Proportions of explained volatility of good-level real exchange rates: Bils and Klenow (2004)

Calvo pricing with various ρ

ρ	0	0.3	0.5	0.7	0.9	0.95	0.98	0.521
Theory/Data	0.130	0.143	0.153	0.148	0.096	0.064	0.036	0.153

Dual stickiness pricing with $\rho = 0.83$

ω	0	0.3	0.5	0.7	0.9	0.95	0.98	0.940
Theory/Data	0.125	0.149	0.181	0.276	0.691	1.129	1.950	1.000

NOTES: Numbers are median ratios of the theoretical volatility, predicted by Bils and Klenow (2004), to observed volatility measured by normalized standard deviation of real exchange rate data. Theoretical volatility for the upper panel is the normalized standard deviation for various ρ when Calvo pricing is used. Theoretical volatility for the lower panel is the normalized standard deviation for various common ω with $\rho = 0.83$ when dual stickiness pricing is used. The normalized sample standard deviation of real exchange rate is the extracted standard deviation component due to time specific shocks in the two-way error component model. The last column of each panel shows the value of ρ and ω, giving the median ratio closest to one.

Table 5: Proportions of explained volatility of good-level real exchange rates: Nakamura and Steinsson (2007)

Calvo pricing with various ρ

ρ	0	0.3	0.5	0.7	0.9	0.95	0.99	0.801
Theory/Data	0.234	0.298	0.351	0.403	0.398	0.312	0.212	0.426

Dual stickiness pricing with $\rho = 0.83$

ω	0	0.3	0.5	0.7	0.9	0.95	0.98	0.916
Theory/Data	0.423	0.449	0.478	0.562	0.882	1.311	2.082	1.000

NOTES: See the notes of Table 4. Nakamura and Steinsson's (2007) data, instead of Bils and Klenow's (2004), is used for the computation of the theoretical value.

Table 6: Intervals between information update

		one month or less	1.01-5.99 months	6-11.99 months	12 months or above
Blinder et. al.'s (1998) survey data		25.6	13.2	16.5	44.5
Bils and Klenow	Persistence	11.5	8.5	26.7	53.3
	Volatility	6.1	4.2	18.2	71.5
Nakamura	Persistence	33.3	12.7	18.2	35.8
and Steinsson	Volatility	21.8	13.9	14.5	49.7

NOTES: The numbers in the first row represent the distribution, in percentages, of the frequency of price reviews reported in Blinder, Canetti, Lebow, and Rudd (1998, Table 4.7 in p. 90). The second and third rows show the distribution of information delay implied by the observed persistence and volatility of real exchange rates based on Bils and Klenow's (2004) data on the frequency of price changes. The fourth and fifth rows show the distribution of information delay when Nakamura and Steinsson's (2007) data on the frequency of regular price changes is used.

Table A1: Frequency of price changes and information updates by category

ELI	Category name	Price Bils & Klenow	Information implied by Per.	Information implied by Vol.	Price Nakamura & Steinsson	Information implied by Per.	Information implied by Vol.	# of goods
FA	Cereals and cereals products	26.5	11.1	4.9	11.5	100.0	6.9	7
FB	Bakery products	25.7	5.5	4.4	9.8	8.7	6.8	1
FC	Beef and veal	47.2	12.2	4.8	25.5	13.8	5.5	8
FD	Pork	47.9	10.2	3.8	23.2	12.3	4.4	6
FF	Poultry	39.4	53.0	2.7	16.6	53.6	3.1	2
FG	Fish and seafood	42.4	8.7	10.6	20.4	9.7	15.2	1
FH	Eggs	61.8	7.5	6.5	47.6	7.5	6.8	1
FJ	Dairy and related products	33.7	6.7	4.4	24.9	7.2	5.3	4
FK	Fresh fruits	36.4	7.5	5.6	16.6	17.3	6.9	8
FL	Fresh vegetables	62.4	24.4	3.4	40.8	25.6	3.6	6
FM	Processed fruits and vegetables	24.9	5.2	4.1	10.5	7.7	6.0	6
FN	Juices and nonalcoholic drinks	35.6	6.1	2.4	13.1	8.2	2.9	4
FP	Beverage incl. coffee and tea	21.1	8.8	7.3	8.9	18.1	13.2	11
FR	Sugar and sweets	22.9	4.8	7.0	9.9	7.1	12.7	2
FS	Fats and oils	29.5	14.5	6.1	18.1	16.0	6.7	8
FV	Food away from home	9.0	3.6	12.9	5.0	5.9	88.5	3
FW	Alcoholic beverages at home	19.3	6.1	6.8	10.6	7.5	10.0	7
FX	Alcoholic beverages away from home	6.4	2.4	14.1	5.0	3.0	25.1	1
HB	Lodging away from home	38.1	11.2	4.9	41.7	11.2	4.8	2
HF	Gas and electricity	43.4	3.6	5.3	38.1	3.6	5.4	1
HK	Appliances	19.0	2.7	3.6	3.6	15.0	25.6	2
HL	Other equipment and furnishings	16.1	10.2	6.6	2.8	100.0	100.0	1
HN	Housekeeping supplies	19.2	9.1	3.2	9.4	60.0	5.7	8
HP	Household operations	6.5	6.7	38.6	4.3	10.8	100.0	1
AA	Men's apparel	26.0	3.1	7.5	4.5	11.3	100.0	5
AB	Boy's apparel	25.9	2.4	11.5	4.3	6.9	100.0	1
AC	Women's apparel	45.0	6.3	6.8	2.5	100.0	100.0	6
AE	Footwear	28.0	4.8	7.1	3.5	60.0	100.0	2
AF	Infants' and toddlers' apparel	36.3	7.6	7.8	3.5	100.0	100.0	2
TA	New and used motor vehicles	39.1	7.5	5.7	31.3	7.6	6.0	7
TB	Motor fuel	78.9	11.3	6.3	88.6	11.3	6.2	1
TD	Motor vehicle maintenance and repair	11.6	6.7	6.1	10.7	7.1	6.4	2
TE	Motor vehicle insurance	15.5	3.2	11.8	8.2	4.6	27.7	1
TG	Public transportation	5.0	4.3	19.8	4.4	4.9	31.2	3
MB	Nonprescription drugs and medical supplies	13.7	5.8	14.8	7.9	8.7	42.6	2
RA	Video and audio	22.0	10.3	10.2	9.4	55.7	24.8	2
RD	Photography	8.6	9.6	16.2	8.8	12.0	30.5	2
RF	Recreation services	8.8	6.7	13.3	9.0	6.6	12.9	1
RG	Recreational reading materials	12.4	15.1	34.5	5.4	100.0	100.0	3
GA	Tobacco and smoking products	21.6	4.3	1.3	23.2	4.3	1.3	4
GB	Personal care products	11.1	4.7	10.8	3.9	14.7	100.0	10
GC	Personal care services	4.1	78.7	100.0	3.1	100.0	100.0	2
GD	Miscellaneous personal services	5.1	13.8	100.0	3.0	100.0	100.0	8

NOTES: ELI in the first column stands for the entry level item in the CPI. EIU price series for good and service used in the analysis are matched to BLS's ELI codes. The third column shows the median value of average monthly frequencies of price changes from Bils and Klenow (2004), among the goods included in each category. The fourth and fifth columns show the median value of the estimated average monthly frequencies of information updates implied by the persistence (Per.) and volatility (Vol.) of good-level real exchange rates, when Bils and Klenow (2004) is used to compute the theoretical prediction. The sixth column is the median of the frequencies of regular price changes from Nakamura and Steinsson (2007). The seventh and eighth columns show the median of frequencies of information updates when Nakamura and Steinsson's (2007) data is used. The last column shows the total numbers of goods and services included in each category of ELI codes.