Estimating the Cross-sectional Distribution of Price Stickiness from Aggregate Data*

Carlos Carvalho
Federal Reserve Bank of New York

Niels Dam
Danmarks Nationalbank

June 2009

*The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve System, or the Danmarks Nationalbank.
Motivation - 1

- “Micro-macro puzzle” in monetary economics
  - Apparent disconnect between microeconomic evidence on price stickiness and behavior of macroeconomic variables
  - Simple sticky-price models require “too much” price rigidity to match aggregate data ("simple" = without many frictions)
  - Disconnect usually documented in “homogeneous firms” models

- Micro data on prices
  - Ample evidence of heterogeneity in price stickiness; ignored in most empirical papers
  - Has been shown to matter in calibrated models
Motivation - 2

- Idea:
  - Ask whether aggregate data provide evidence for such type of heterogeneity
  - Does this help get rid of the “micro-macro disconnect”?

- Issue: informational content of aggregates with respect to distribution of price stickiness

- “Non-field-specific motivation”
  - Cross-sectional inference using only aggregate data
  - Remark: of course requires *some* structure (i.e., inference given a model)
This Paper - 1

- “Semi-structural” multi-sector sticky-price model
  - Supply-side: multi-sector Taylor-staggered-contracts economy
  - Rest of the model: exogenous time-series processes ("restrictions in a Var")

- Are aggregate data informative about cross-sectional distribution of price stickiness?
  - Aggregate data = \{nominal GDP, real GDP\}
• If so, how does estimated distribution of stickiness compare with micro data?

• Can we discriminate between homogeneous- and heterogeneous-firms models?

• Estimation with Bayesian methods
  – “Flat” priors (no micro data)
  – Informative priors (using micro data)
Findings - 1

- Aggregate data are informative of cross-sectional distribution of price stickiness
  - Intuition: “frequency composition effect” (Carvalho 2006)
  - Monte Carlo exercise: small vs. large sample

- Empirical results with flat priors:
  - Distributions resemble those derived from micro data
  - *All* sectors with < 1 year of price rigidity: quite close to Bils and Klenow (2004); but incredible degree of pricing complementarity
  - *Some* sectors with up to 1.5 or 2 years of rigidity: looks similar to Nakamura and Steinsson (2008); more reasonable complementarities
  - Strong evidence in favor of latter specifications
Findings - 2

- Empirical results with informative priors:
  - (Dirichlet) Priors on cross-sectional distribution of stickiness derived from micro data
  - Bils-Klenow, Nakamura-Steinsson priors
  - Strong evidence in favor of models with *some* prices lasting up to 1.5 - 2 years
  - Remarkably stable results
Homogeneous (1-sector) vs. Heterogeneous (multi-sector) models:

- Strong support for heterogeneity
- Lowest posterior odds ratio is of the order of $10^6 : 1$
- Best 1-sector model: all prices last for 7 quarters
Outline

- Overview of the model

- Identification: intuition and Monte Carlo evidence

- Empirical results

- Robustness + more on identification

- Conclusion
Model - 1

- Continuum of monopolistically competitive firms
  - Each firm produces unique variety, faces a demand that depends negatively on its relative price
  - $K$ sectors that differ (only) in the frequency of price changes
  - Distribution of firms: $(\omega_1, \ldots, \omega_K)$ with $\omega_k > 0$, $\sum_{k=1}^{K} \omega_k = 1$; $\omega_k$'s give mass of firms in sector $k$
  - Taylor staggered price setting: firms in sector $k$ set prices for $k$ periods
Model - 2

- Optimal prices chosen: $X_t(k, j)$

$$\max \mathcal{E} \sum_{i=0}^{k-1} Q_{t,t+i} \prod \left( X_t(k, j), P_{t+i}, Y_{t+i}, \xi_{t+i} \right),$$

- Prices charged:
  - $\{P_t(k, j)\}_{j \in [0,1]}$
  - Note: $P_{t+k-1}(k, j) = P_{t+k-2}(k, j) = \ldots = P_t(k, j) = X_t(k, j) = X_t(k)$
Model - 3

- Sectoral prices:

\[ P_t(k) = \Lambda \left( \{ P_t(k, j) \}_{j \in [0, 1]} \right) = \Lambda \left( \{ X_{t-i}(k) \}_{i=0, \ldots, k-1} \right) \]

- Aggregate price:

\[ P_t = \Gamma \left( \{ P_t(k), \omega_k \}_{k=1, \ldots, K} \right) \]

- Assume deterministic zero-inflation steady state:

\[ \xi_{t+i} = 0, Y_{t+i} = \bar{Y}, Q_{t, t+i} = \beta, \text{ and for all } (k, j), \ X_t(k, j) = P_t = \bar{P} \]
Model - 4

- Loglinear model:

\[
 x_t(k) = \frac{1 - \beta}{1 - \beta^k} E_t \sum_{i=0}^{k-1} \beta^i \left( p_{t+i} + \zeta \left( y_{t+i} - y^n_{t+i} \right) \right)
\]

\[
 p_t = \sum_{k=1}^{K} \omega_k p_t(k)
\]

\[
 p_t(k) = \int_0^1 p_t(k, j) \, dj = \frac{1}{k} \sum_{j=0}^{k-1} x_{t-j}(k)
\]

\[
 p_t + y_t = m_t = \rho_0 + \rho_1 m_{t-1} + \ldots + \rho_p m_{t-p} + \varepsilon^m_t
\]

\[
 y^n_t = \delta_0 + \delta_1 y^n_{t-1} + \ldots + \delta_p y^n_{t-p} + \varepsilon^n_t
\]
Monte Carlo Preliminaries

- Artificial data on nominal and real GDP
- MLE estimation
- Large sample (75 samples of 1000 quarters); small sample (240 samples of 100 quarters)
- Informational content of the aggregate data
Table 1: Monte-Carlo - Maximum Likelihood estimation

<table>
<thead>
<tr>
<th>ζ</th>
<th>True</th>
<th>Mean</th>
<th>5\textsuperscript{th} perc.</th>
<th>95\textsuperscript{th} perc.</th>
<th>Mean</th>
<th>5\textsuperscript{th} perc.</th>
<th>95\textsuperscript{th} perc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.106</td>
<td>0.059</td>
<td>0.15</td>
<td>0.179</td>
<td>0.022</td>
<td>0.415</td>
<td></td>
</tr>
<tr>
<td>ω₁</td>
<td>0.40</td>
<td>0.395</td>
<td>0.183</td>
<td>0.621</td>
<td>0.318</td>
<td>0.033</td>
<td>0.871</td>
</tr>
<tr>
<td>ω₂</td>
<td>0.10</td>
<td>0.100</td>
<td>0.000</td>
<td>0.257</td>
<td>0.096</td>
<td>0.000</td>
<td>0.376</td>
</tr>
<tr>
<td>ω₃</td>
<td>0.10</td>
<td>0.091</td>
<td>0.000</td>
<td>0.197</td>
<td>0.088</td>
<td>0.000</td>
<td>0.304</td>
</tr>
<tr>
<td>ω₄</td>
<td>0.40</td>
<td>0.414</td>
<td>0.233</td>
<td>0.570</td>
<td>0.498</td>
<td>0.064</td>
<td>0.801</td>
</tr>
</tbody>
</table>
Estimation - observables

- Observables: nominal and real GDP (linear detrending), 1983-2007 (pre-sample of 16 quarters)

Real GDP ("output")

Nominal GDP ("nominal income")
Estimation - priors 1

- Loose priors for parameters other than $\omega$'s

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
<th>Mode</th>
<th>Mean</th>
<th>Std.dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\zeta$</td>
<td>Gamma(1.2, 0.2)</td>
<td>1.00</td>
<td>6.00</td>
<td>5.48</td>
</tr>
<tr>
<td>$\sigma_n, \sigma_m$</td>
<td>Gamma(1.5, 20)</td>
<td>0.025</td>
<td>0.075</td>
<td>0.06</td>
</tr>
<tr>
<td>$\rho_j, \delta_j$</td>
<td>N(0, $5^2$)</td>
<td>0.00</td>
<td>0.00</td>
<td>5.00</td>
</tr>
</tbody>
</table>
Flat and informative priors for $\omega$’s

- Dirichlet prior on $\omega$’s

$$f_\omega (\omega | \alpha) \propto \prod_{k=1}^{K} \omega_k^{\alpha_k-1}, \quad \forall \alpha_k > 0, \forall \omega_k \geq 0, \quad \sum_{k=1}^{K} \omega_k = 1$$

- $\alpha_0 \equiv \sum_k \alpha_k$ gives “overall tightness”

- Set $\alpha_k = 1 + \hat{\omega}_k (\alpha_0 - K)$, where $\hat{\omega}_k$ are empirical weights from BK or NS

- $\alpha_k > 1$: mode is set to $(\hat{\omega}_1, \ldots, \hat{\omega}_K)$

- $\alpha_0 = K$: flat prior
Table 2: Cross-sectional distributions of price stickiness

<table>
<thead>
<tr>
<th>Parameter</th>
<th>BK</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\omega}_1$</td>
<td>0.40</td>
<td>0.27</td>
</tr>
<tr>
<td>$\hat{\omega}_2$</td>
<td>0.24</td>
<td>0.07</td>
</tr>
<tr>
<td>$\hat{\omega}_3$</td>
<td>0.12</td>
<td>0.10</td>
</tr>
<tr>
<td>$\hat{\omega}_4$</td>
<td>0.12</td>
<td>0.11</td>
</tr>
<tr>
<td>$\hat{\omega}_5$</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>$\hat{\omega}_6$</td>
<td>0.03</td>
<td>0.13</td>
</tr>
<tr>
<td>$\hat{\omega}_7$</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>$\hat{\omega}_8$</td>
<td>0.03</td>
<td>0.20</td>
</tr>
<tr>
<td>$\hat{\sigma}_k(*)$</td>
<td>2.54</td>
<td>4.23</td>
</tr>
<tr>
<td>$\hat{\sigma}_k(*)$</td>
<td>1.86</td>
<td>2.66</td>
</tr>
</tbody>
</table>
Estimation - details

- Bayesian approach (Random-Walk Metropolis algorithm)

- Initial maximization of posterior

- Transform all parameters to full support on the real line (important to handle rejection rates)

- Run three successive adaptive phases of 900,000 iterations in total
  - Adjust and fine-tune jumping covariance matrix

- Run fixed phase: 5 parallel chains of 300,000 each
  - Check for convergence and combine last 2/3’s for simulated sample of 1 million draws
Empirical results - flat priors, $K=4$

- $K = 4$
Empirical results - flat priors, $K=6$

- $K = 6$
Empirical results - flat priors, $K=8$

- $K = 8$
Empirical results - "micro vs. macro estimates"

- Flat priors - summary

<table>
<thead>
<tr>
<th>$K$</th>
<th>$Corr \rightarrow BK$</th>
<th>$Corr \rightarrow NS$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>96.73%</td>
<td>31.31%</td>
</tr>
<tr>
<td>6</td>
<td>56.00%</td>
<td>72.64%</td>
</tr>
<tr>
<td>8</td>
<td>43.04%</td>
<td>62.79%</td>
</tr>
</tbody>
</table>
Figure 10: Flat prior (dashed line) and posterior (solid line) marginal distributions, $K = 6$
## Estimation Summary

<table>
<thead>
<tr>
<th>Prior</th>
<th>$K = 4$</th>
<th>$K = 6$</th>
<th>$K = 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Flat</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \zeta )</td>
<td>4.440</td>
<td>4.440</td>
<td>4.440</td>
</tr>
<tr>
<td>( (0.47;16.86) )</td>
<td>(0.00;0.02)</td>
<td>(0.01;0.06)</td>
<td>(0.01;0.11)</td>
</tr>
<tr>
<td>( \bar{k} )</td>
<td>2.499</td>
<td>3.504</td>
<td>4.503</td>
</tr>
<tr>
<td>( (1.67;3.32) )</td>
<td>(1.39;2.48)</td>
<td>(2.44;4.56)</td>
<td>(3.25;5.76)</td>
</tr>
<tr>
<td>( \sigma_k )</td>
<td>0.983</td>
<td>1.559</td>
<td>2.137</td>
</tr>
<tr>
<td>( (0.02;1.32) )</td>
<td>(0.88;1.39)</td>
<td>(1.09;2.01)</td>
<td>(1.58;2.68)</td>
</tr>
<tr>
<td>log MD</td>
<td>794.993</td>
<td>806.418</td>
<td>808.031</td>
</tr>
<tr>
<td><strong>Loose NS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \zeta )</td>
<td>4.440</td>
<td>4.440</td>
<td>4.440</td>
</tr>
<tr>
<td>( (0.47;16.86) )</td>
<td>(0.00;0.02)</td>
<td>(0.01;0.06)</td>
<td>(0.02;0.11)</td>
</tr>
<tr>
<td>( \bar{k} )</td>
<td>2.736</td>
<td>3.637</td>
<td>4.364</td>
</tr>
<tr>
<td>( (2.02;3.38) )</td>
<td>(1.64;2.61)</td>
<td>(2.75;4.51)</td>
<td>(3.38;5.96)</td>
</tr>
<tr>
<td>( \sigma_k )</td>
<td>1.172</td>
<td>1.843</td>
<td>2.405</td>
</tr>
<tr>
<td>( (0.08;1.38) )</td>
<td>(1.07;1.42)</td>
<td>(1.49;2.15)</td>
<td>(1.96;2.79)</td>
</tr>
<tr>
<td>log MD</td>
<td>795.011</td>
<td>806.944</td>
<td>808.270</td>
</tr>
<tr>
<td><strong>Loose BK</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \zeta )</td>
<td>4.440</td>
<td>4.440</td>
<td>4.440</td>
</tr>
<tr>
<td>( (0.47;16.86) )</td>
<td>(0.00;0.02)</td>
<td>(0.01;0.05)</td>
<td>(0.01;0.08)</td>
</tr>
<tr>
<td>( \bar{k} )</td>
<td>2.350</td>
<td>2.951</td>
<td>3.495</td>
</tr>
<tr>
<td>( (1.75;3.02) )</td>
<td>(1.54;2.49)</td>
<td>(2.21;3.80)</td>
<td>(2.64;4.49)</td>
</tr>
<tr>
<td>( \sigma_k )</td>
<td>1.104</td>
<td>1.668</td>
<td>2.222</td>
</tr>
<tr>
<td>( (0.02;1.33) )</td>
<td>(0.97;1.37)</td>
<td>(1.29;2.02)</td>
<td>(1.75;2.67)</td>
</tr>
<tr>
<td>log MD</td>
<td>795.350</td>
<td>806.024</td>
<td>807.563</td>
</tr>
</tbody>
</table>
Dynamics - Het. vs. Hom. models

Response of $Y_t$ to $e_{t}^{m}$

Response of $\pi_t$ to $e_{t}^{m}$

Response of $Y_t$ to $e_{t}^{n}$

Response of $\pi_t$ to $e_{t}^{n}$
Summary of findings

- Estimated cross-sectional distribution resembles BK priors for $K = 4$, but NS priors for $K = 6, 8$

- Allowing for $K > 4$ is important: least posterior odds against $K = 4$ are of the order of $10^3 : 1$ to $10^5 : 1$

- Higher nominal rigidities imply lower real rigidity (strategic complementarity)

- For $K = 4$, incredible degree of real rigidity (matches Coenen et al. 2007)

- Heterogeneity empirically important: lowest posterior odds against homogeneous models is of the order of $10^6 : 1$
More on identification

- Identification depends critically on relative variance of observable versus unobservable driving processes
  - More difficult when unobservable process drives most of the dynamics

- Extreme case
  - Two sector economy: 1 sticky-price sector + 1 flex-price sector
  - Output is observed; one driving process:
    - If driving process is observed: have identification
    - If driving process is unobserved: have observational equivalence
Robustness - I

- Results robust to different priors, different specifications for exogenous processes (e.g. AR(3))

- Not all results hold with Calvo pricing:
  - Identification is “harder” with Calvo pricing
  - Intuition: in 1-sector models, different frequencies of price changes confounded with different degrees of real rigidity
  - As a result: differences in aggregate implications of various frequencies of price changes are not as stark
Robustness - II

- But strictly speaking we have identification

- Origin of problem in practice: high variance of unobservable exogenous process

- Fully-specified DSGE + more observables might help: implicit restrictions on unobservable processes

- Nevertheless: can still discriminate between 1-sector and multi-sector models
  - Likelihood ratio test rejects 1-sector in favor of 2-sector model
Conclusions

- Micro-macro consistency

- Strong evidence for heterogeneity

- “Enough” heterogeneity is important

- Informational content of aggregate data
  
  - Other settings in which this inference is possible?
  
  - Estimation for countries/periods for which no micro data is available