Inflation at the Household Level

Greg Kaplan, University of Chicago and NBER
Sam Schulhofer-Wohl, Federal Reserve Bank of Chicago

San Francisco Fed Conference on Macroeconomics and Monetary Policy, March 31, 2017

The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System.
Overview

- We estimate inflation rates household by household.

- Lots of heterogeneity: interquartile range of annual rates varies between 6.2 and 9.0 percentage points.
  - When aggregate inflation is 2%, at least half of households have inflation rates above 5% or below −1%.

- Sources of heterogeneity:
  - Different weights on broad consumption categories. (small)
  - Different product choices within categories. (big, new)
  - Different prices for identical products. (big, new)

- Aggregate inflation accounts for less than 1/10 of variability in household inflation over time.
Outline

• Data and inflation calculations.

• Cross-sectional properties of household inflation.

• Time-series properties of household inflation.
Data and calculations
Data

- Panel of 50,000 households, replenished for attrition.
- Consumption of **goods with barcodes**.
- Each shopping trip (including online shopping): record barcode and price of each item purchased.
  - Price = average at store for the week if Nielsen covers store.
  - Otherwise, household records price.
Estimating household inflation rates

- Need to
  - Define household-level consumption bundles.
  - Measure change in *household’s price* for each good between two dates ⇒ must see household buy good at both dates.

- Aggregate purchases to quarterly frequency.

- To remove (most) seasonality, compare two quarters one year apart.

- Exclude product if household’s price changes by factor > 3.

- Exclude households with < 5 matched barcodes.
  - 77% of HH that buy something at $t$ also buy something at $t + 4$, 72% at least 5 matched barcodes.
## Distribution of spending (%)

<table>
<thead>
<tr>
<th>Category</th>
<th>CPI-U</th>
<th>all spending</th>
<th>5+ matched barcodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and beverages</td>
<td>15.26</td>
<td>61.22</td>
<td>74.38</td>
</tr>
<tr>
<td>Food</td>
<td>14.31</td>
<td>58.08</td>
<td>67.61</td>
</tr>
<tr>
<td>Food at home</td>
<td>8.60</td>
<td>53.87</td>
<td>64.77</td>
</tr>
<tr>
<td>Food away from home</td>
<td>5.71</td>
<td>4.22</td>
<td>2.83</td>
</tr>
<tr>
<td>Alcoholic beverages</td>
<td>0.95</td>
<td>3.13</td>
<td>6.78</td>
</tr>
<tr>
<td>Housing</td>
<td>41.02</td>
<td>9.03</td>
<td>5.11</td>
</tr>
<tr>
<td>Apparel</td>
<td>3.56</td>
<td>8.40</td>
<td>-</td>
</tr>
<tr>
<td>Transportation</td>
<td>16.85</td>
<td>0.22</td>
<td>0.14</td>
</tr>
<tr>
<td>Medical care</td>
<td>7.16</td>
<td>6.92</td>
<td>4.85</td>
</tr>
<tr>
<td>Recreation</td>
<td>5.99</td>
<td>6.57</td>
<td>5.85</td>
</tr>
<tr>
<td>Education and communication</td>
<td>6.78</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>3.38</td>
<td>7.64</td>
<td>9.67</td>
</tr>
<tr>
<td>Tobacco and smoking products</td>
<td>0.81</td>
<td>1.87</td>
<td>6.46</td>
</tr>
</tbody>
</table>
Inflation rates with CPI vs. KNCP bundles

All indexes use CPI prices.
Four ways to construct household inflation indexes

- **Household-level prices:**
  - Household’s consumption bundle at barcode level.
  - Household’s price paid for each barcode.

- **Barcode-average prices:**
  - Household’s consumption bundle at barcode level.
  - National average price paid for each barcode.

- **CPI prices:**
  - Household’s consumption bundle at level of broad categories.
  - Item stratum price indexes from CPI.
  - Comparable to previous literature.

- **Shopping-trip prices (in progress):**
  - Price when household shopped, whether it bought UPC or not.
Household inflation indexes

- Notation: household $i$, UPC $j$, date $t$.

- Laspeyres with household-level prices:

  \[ \pi_{it,t+4}^L = \frac{\sum_{j: q_{ijt}, q_{ij}, t+4 > 0} q_{ijt} p_{ij,t+4}}{\sum_{j: q_{ijt}, q_{ij}, t+4 > 0} q_{ijt} p_{ijt}} \]

- Laspeyres with barcode-average prices:

  \[ \pi_{it,t+4}^{L,BC} = \frac{\sum_{j: q_{ijt}, q_{ij}, t+4 > 0} q_{ijt} \bar{p}_j, t+4}{\sum_{j: q_{ijt}, q_{ij}, t+4 > 0} q_{ijt} \bar{p}_jt} \]

- Laspeyres with CPI prices:

  \[ \pi_{it,t+4}^{L,CPI} = \sum_{j: q_{ijt}, q_{ij}, t+4 > 0} s_{ij,t+4}^L \left( \frac{p_{k(j),t+4}^{CPI}}{p_{k(j),t}^{CPI}} \right) \]

  $k(j)$: CPI item stratum, $s_{ij}^L$: initial budget share of UPC $j$. 
Cross-sectional properties of household inflation
Inflation distribution, 2004q4–2005q4

The graph shows the distribution of household inflation rates from 2004q4 to 2005q4. The x-axis represents the household inflation rate (%), and the y-axis shows the density. The graph compares three types of prices:

- **Household-level prices** (blue line)
- **Barcode-average prices** (orange line)
- **CPI prices** (black line)

The data indicates a higher concentration of inflation rates around 0%, with slight variations for other rates.
Interquartile range of inflation rates

- Household-level prices
- Barcode-average prices
- CPI prices
Bundles with few UPCs don’t drive dispersion
Evolution of the inflation distribution

![Graph showing the evolution of inflation distribution from 2004 to 2012. The graph includes lines for 10th-90th percentiles, median, mean, 25th-75th percentiles, and aggregate index.](image-url)
Low-income households usually have higher inflation rates.
How much heterogeneity do demographics explain?

- OLS and quantile regressions of $\pi_{it,t+4}^L - \pi_{t,t+4}^{L,CPI}$ on large vector of demographics:
  - Household income
  - Age of head(s)
  - Education of head(s)
  - Region
  - Household size and composition
  - Race
- Control for time dummies.
- 835,386 household-quarter observations.
- Most variance remains unexplained:
  - OLS $R^2$, time dummies only: 0.009
  - OLS $R^2$, time dummies plus all demographics: 0.012
## Inflation vs. income and education

<table>
<thead>
<tr>
<th></th>
<th>(1) OLS</th>
<th></th>
<th>(2) Median</th>
<th></th>
<th>(3) IQR</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coeff.</td>
<td>std. err.</td>
<td>coeff.</td>
<td>std. err.</td>
<td>coeff.</td>
<td>std. err.</td>
</tr>
<tr>
<td><strong>household income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$20,000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$20,000–$39,999</td>
<td>-0.206</td>
<td>(0.055)</td>
<td>-0.126</td>
<td>(0.039)</td>
<td>-0.399</td>
<td>(0.079)</td>
</tr>
<tr>
<td>$40,000–$59,999</td>
<td>-0.420</td>
<td>(0.052)</td>
<td>-0.257</td>
<td>(0.041)</td>
<td>-0.597</td>
<td>(0.085)</td>
</tr>
<tr>
<td>$60,000–$99,999</td>
<td>-0.587</td>
<td>(0.059)</td>
<td>-0.468</td>
<td>(0.045)</td>
<td>-0.706</td>
<td>(0.086)</td>
</tr>
<tr>
<td>$100,000</td>
<td>-0.731</td>
<td>(0.065)</td>
<td>-0.597</td>
<td>(0.050)</td>
<td>-0.873</td>
<td>(0.096)</td>
</tr>
<tr>
<td><strong>highest education of household head(s)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; high school</td>
<td>-0.064</td>
<td>(0.127)</td>
<td>-0.029</td>
<td>(0.108)</td>
<td>-0.280</td>
<td>(0.167)</td>
</tr>
<tr>
<td>high school diploma</td>
<td>-0.138</td>
<td>(0.127)</td>
<td>-0.102</td>
<td>(0.107)</td>
<td>-0.118</td>
<td>(0.165)</td>
</tr>
<tr>
<td>some college</td>
<td>-0.251</td>
<td>(0.128)</td>
<td>-0.163</td>
<td>(0.110)</td>
<td>-0.099</td>
<td>(0.180)</td>
</tr>
<tr>
<td>bachelor’s degree</td>
<td>-0.285</td>
<td>(0.139)</td>
<td>-0.137</td>
<td>(0.118)</td>
<td>0.024</td>
<td>(0.185)</td>
</tr>
</tbody>
</table>

Other controls: age, region, HH size/composition, race, time dummies
Search theory: bargain hunting

- Equilibrium models of search and price dispersion hold that households pay lower prices when they observe more prices.

<table>
<thead>
<tr>
<th></th>
<th>household $\pi$ - aggregate $\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) Median</td>
</tr>
<tr>
<td></td>
<td>coeff.</td>
</tr>
<tr>
<td>log(# of shopping trips)</td>
<td></td>
</tr>
<tr>
<td>initial quarter</td>
<td>0.352</td>
</tr>
<tr>
<td>final quarter</td>
<td>-0.409</td>
</tr>
</tbody>
</table>

+ demographic controls & time dummies

- Coefficients imply that households who make more shopping trips pay lower prices and have less-dispersed inflation rates.
Demand theory: substitution between goods

- As prices change, households should substitute toward goods whose relative prices have fallen.

- Implies $\pi^L > \pi^F > \pi^P$ because Laspeyres uses initial-period bundle and Paasche uses final-period bundle.

- $\pi^L - \pi^F = \text{substitution bias}$.
  - Boskin Commission: 0.4 percentage point in aggregate CPI.

- What are the substitution patterns in KNCP data?
Mean differences from Fisher index

Laspeyres Paasche
Laspeyres vs. Paasche inflation rates, 2004q4–2005q4

density

0.05 0.1 0.15 0.2

−10 −5 0 5 10 15
difference between household Laspeyres and Paasche inflation rates (%)
Intertemporal choice: do households buy more when they face a lower price level?

• Growth rate of spending:

\[ \ln x_{i,t+4} - \ln x_{it} = \ln \pi_{it,t+4} + \ln q_{i,t+4} - \ln q_{it} \]

⇒ recover quantity index \( \Delta \ln q \) given spending \( x \) and inflation \( \pi \).

• Variance decomposition (on average across quarters):

\[
\begin{align*}
\text{Var}(\ln \pi_{it,t+4}) & \quad 0.007 \\
\text{+ Var}(\ln q_{i,t+4} - \ln q_{it}) & \quad 0.113 \\
\text{+ 2Cov}(\ln_{it,t+4}, \ln q_{i,t+4} - \ln q_{it}) & \quad -0.004 \\
\hline
= \text{Var}(\ln x_{i,t+4} - \ln x_{it}) & \quad 0.116
\end{align*}
\]

• In a structural model, could recover EIS from this covariance.
Time-series properties of household inflation
Distribution of 1- and 2-year inflation rates

(a) household prices
(b) barcode–average prices
(c) CPI prices

Standard deviations of inflation rates

(a) 1–year
(b) 2–year

- household–level prices
- barcode–average prices
- CPI prices
Serial correlation of 1-year inflation rates

household−level prices   barcode−average prices   CPI prices
A simple model of the stochastic process

- Log deviation of HH price level from aggregate: FE plus AR(1)

\[
\log P_{it} - \log P_t = \mu_i + z_{it}
\]

\[
z_{it} = \rho z_{i,t-4} + \epsilon_{it}
\]

- Assume initial conditions from ergodic distribution. Then

\[
\rho = 1 + 2\text{Corr}(\pi_{it}, \pi_{i,t-1})
\]

\[
\text{Corr}(\pi_{it}, \pi_{i,t-1}) = -0.1 \quad \Rightarrow \quad \rho = 0.8
\]

- Variance decomposition of \(\pi_{it}\):
  - Cross-sectional s.d. of \(\pi_{it}\): 6.2 percentage points.
  - Time-series s.d. of aggregate \(\pi\): 1.9 percentage points.
  - \(\Rightarrow\) 91% of variance of \(\pi_{it}\) comes from heterogeneity.
Conclusion
Implications

- Household inflation rates are highly heterogeneous.

- Household price levels deviate persistently from aggregate price level.

- Could use results to calibrate models of individual inference about aggregate inflation.
  - Shopping-trip prices may be helpful here.

- Challenges for monetary economics:
  - Welfare with heterogeneous inflation rates?
  - Heterogeneous real interest rates for given nominal interest rate?
  - How well can households forecast their own inflation rates?
Inflation at the Household Level

Greg Kaplan, University of Chicago and NBER
Sam Schulhofer-Wohl, Federal Reserve Bank of Chicago

San Francisco Fed Conference on Macroeconomics and Monetary Policy, March 31, 2017

The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System.
Distribution of deviations between KNCP and CPI stratum inflation rates

![Graph showing distribution of deviations between KNCP and CPI stratum inflation rates.](Image)
Aggregate inflation rates computed with KNCP and CPI stratum prices

Both indexes use KNCP bundle.
Share of variance from common prices

![Graph showing the share of variance from common prices over the years 2004 to 2012. The graph compares barcode-average prices with CPI prices, highlighting the variance fluctuations over time.]
## Interquartile range with different indexes

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>s.d.</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household-level prices:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laspeyres</td>
<td>7.33</td>
<td>0.74</td>
<td>6.23</td>
<td>8.99</td>
</tr>
<tr>
<td>Fisher</td>
<td>7.13</td>
<td>0.72</td>
<td>6.12</td>
<td>8.92</td>
</tr>
<tr>
<td>Paasche</td>
<td>7.37</td>
<td>0.76</td>
<td>6.34</td>
<td>9.18</td>
</tr>
<tr>
<td><strong>Barcode-average prices:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laspeyres</td>
<td>3.99</td>
<td>0.77</td>
<td>3.06</td>
<td>5.73</td>
</tr>
<tr>
<td>Fisher</td>
<td>3.87</td>
<td>0.75</td>
<td>2.95</td>
<td>5.68</td>
</tr>
<tr>
<td>Paasche</td>
<td>3.98</td>
<td>0.76</td>
<td>3.03</td>
<td>5.81</td>
</tr>
<tr>
<td><strong>Stratum-average prices:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laspeyres</td>
<td>1.96</td>
<td>0.95</td>
<td>0.89</td>
<td>3.96</td>
</tr>
<tr>
<td>Fisher</td>
<td>1.83</td>
<td>0.88</td>
<td>0.91</td>
<td>3.84</td>
</tr>
<tr>
<td>Paasche</td>
<td>1.95</td>
<td>0.92</td>
<td>0.92</td>
<td>3.92</td>
</tr>
<tr>
<td><strong>CPI prices:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laspeyres</td>
<td>1.61</td>
<td>0.80</td>
<td>0.71</td>
<td>3.77</td>
</tr>
<tr>
<td>Fisher</td>
<td>1.57</td>
<td>0.77</td>
<td>0.70</td>
<td>3.53</td>
</tr>
<tr>
<td>Paasche</td>
<td>1.62</td>
<td>0.78</td>
<td>0.71</td>
<td>3.42</td>
</tr>
</tbody>
</table>

Averages from 2004q1 through 2012q3 of IQR for each date.
Quantile regression of household inflation on aggregate inflation

<table>
<thead>
<tr>
<th>Decile</th>
<th>Coefficient on aggregate index</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.011 (0.015)</td>
<td>−7.602 (0.058)</td>
</tr>
<tr>
<td>2</td>
<td>1.013 (0.009)</td>
<td>−4.609 (0.039)</td>
</tr>
<tr>
<td>3</td>
<td>1.026 (0.008)</td>
<td>−2.810 (0.031)</td>
</tr>
<tr>
<td>4</td>
<td>1.052 (0.008)</td>
<td>−1.448 (0.027)</td>
</tr>
<tr>
<td>5</td>
<td>1.093 (0.007)</td>
<td>−0.264 (0.026)</td>
</tr>
<tr>
<td>6</td>
<td>1.137 (0.009)</td>
<td>0.944 (0.030)</td>
</tr>
<tr>
<td>7</td>
<td>1.198 (0.010)</td>
<td>2.286 (0.034)</td>
</tr>
<tr>
<td>8</td>
<td>1.243 (0.012)</td>
<td>4.189 (0.046)</td>
</tr>
<tr>
<td>9</td>
<td>1.305 (0.019)</td>
<td>7.491 (0.066)</td>
</tr>
</tbody>
</table>

835,386 household-quarter observations. Bootstrap standard errors in parentheses.
Low-income households have higher inflation

<table>
<thead>
<tr>
<th>Household income</th>
<th>cumulative inflation (%)</th>
<th>fraction of population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2013q1</td>
<td>2013q2</td>
</tr>
<tr>
<td>$&lt;20,000$</td>
<td>34.35</td>
<td>33.25</td>
</tr>
<tr>
<td></td>
<td>(0.90)</td>
<td>(0.66)</td>
</tr>
<tr>
<td>$20,000–$39,999$</td>
<td>32.37</td>
<td>31.11</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(0.57)</td>
</tr>
<tr>
<td>$40,000–$59,999$</td>
<td>29.90</td>
<td>28.26</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.63)</td>
</tr>
<tr>
<td>$60,000–$99,999$</td>
<td>27.84</td>
<td>25.86</td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.56)</td>
</tr>
<tr>
<td>$\geq$100,000$</td>
<td>25.74</td>
<td>24.23</td>
</tr>
<tr>
<td></td>
<td>(0.65)</td>
<td>(0.71)</td>
</tr>
</tbody>
</table>

Calculated with Laspeyres indexes and household-level prices. Bootstrap standard errors in parentheses.