Trade, Consumer Prices, and Real Income: An Engels Curve Approach

Farrokhi, Jinkins, and Xiang

February 20, 2020

Preliminary: please do not cite without permission

Abstract

We construct consumer price indexes using international trade data and theory inspired by recent quantitative trade research. We control for domestic prices using sectoral level import shares. We compare these trade-inspired consumer price indexes for the United States and Denmark to official consumer price indexes using an approach based on an assumption of constant Engel curves for food due to Hamilton (2001). We find that official consumer price indexes underestimate real income growth in the United States from 1995 to 2015. Our trade-inspired consumer price index correctly reflects real income growth in the United States from 1995-2006, and then underestimates further real income growth from 2007-2015. Our results for Denmark are in progress.

One of the central questions in research in international trade is how much individuals within countries gain from trade between countries. In particular, recent quantitative general equilibrium trade models have been widely deployed to measure these gains from trade (Eaton and Kortum (2002), Costinot and Rodriguez-Clare (2014)). Other researchers have rather focused on disaggregated unit values from customs data or observed prices from scanner data, and used theory to aggregate changes in price to measure how trade affects welfare. Although vast amounts of ink have been spilled on the topic of gains from trade, there is still little consensus. Measured gains are sensitive to the exact specifications of the models employed.

A classic and related problem in empirical macroeconomics is measuring how price indexes change over time. If we do not correctly measure changes in prices over time, we will misestimate growth in production, consumption, and ultimately welfare. There are several well-known challenges to measuring price changes. Among these challenges is that even if goods seem similar, their quality may increase. Another is that disaggreggated prices must be aggregated to the cost of living through assumptions about consumer preferences. Hamilton

(2001) suggests a method to shed light on how closely changes in aggregated price indexes we construct reflect true changes in the price index by exploiting an empirical regularity in consumer behavior. As people become richer, they spend a smaller share of their overall spending on food at home.

The relationship between the share of household expenditures spent on a particular type of good and log total household expenditures is known as an Engel curve. One of the most established empirical regularities in economics data is that the Engel curve for food at home is negative (Chai and Moneta, 2010). More specifically, in a cross-sectional regression of expenditure share on food on the log of total expenditures, studies typically find a slope of around -0.1. This is true for both developing and developed countries, and has remained more or less constant over time. The basic idea of the Hamilton method for evaluating price indexes is to assume that the cross-sectional semi-logarithmic Engel curve for food described above is unchanging over time. Under the assumption of a constant Engel curve for food, including year dummies in a regression of expenditure shares on food on log deflated total expenditures should not change the regression coefficient, and should yield year dummies close to zero. If year dummies are rather statistically significant and large, it suggests that our price index is mismeasured. Moreover, the sign of year dummies suggest the direction of the bias.

One method for deriving a price index consistent with the Engel curve is simply to invert the relationship between nominal income and share of food at home. This method, however, requires data on household expenditures and household characteristics at a fine level. National statistical agencies often either do not have this data, or do not use it in the construction of consumer price indexes. In this paper, we construct a price indexes suggested by recent international trade research, as well as those constructed by national statistics offices. The price indexes we construct can (nearly) all be computed with widely and costlessly available trade and production data.[1] As is standard in the literature following Feenstra (1994), our trade-based index will include a correction for new varieties. In both American and Danish data, we evaluate these price indexes against each other using the Hamilton method described above. We focus on these two countries, because they highlight how differences in trade volumes affect the importance of corrections for the variety margin in international trade. The United States imported only 15% of its GDP in 2018, while Denmark imported 49% of its GDP.[2]. We are in the process of applying for household level data from a larger set of EU countries as well.

Using this method on US data, we find that deflating with CPI consistently underestimates real income growth in the period 1996-2013. The Hamilton method does not reject that deflating income with a consumer price index constructed using only trade unit values and import shares correctly recovers real income in the years from 1996 to 2005. After this period, deflating by the trade based consumer price understates real income just as standard CPI. Danish

[1] We will use a sufficient statistics approach to recover domestic prices.

[2] https://data.worldbank.org/indicator/NE.IMP.GNFS.ZS?year_high_desc=true

results are in progress.

In addition to a theoretical literature connecting Engel curves to the theory on consumer demand (Deaton and Muellbauer, 1980; Banks et al., 1997), there is a more empirical literature using Engel curves to compare welfare across households, time, and locations (Deaton and Muellbauer, 1986; Hamilton, 2001; Almås, 2012; Nakamura et al., 2016). We contribute to this literature by using Engel curves to evaluate price indexes based on results from the international trade literature. Since we take observed trade unit values as the basis for our price index calculations, we build heavily on results from Feenstra (1994); Broda and Weinstein (2006); Redding and Weinstein (2018). Since we ultimately are interested in changes in the price index, our project is also more distantly related to the vast quantitative literature on gains from trade (Eaton and Kortum, 2002; Arkolakis et al., 2012; Bai and Stumpner, 2019), and especially a newer quantitative literature which focuses on gains from trade when preferences are not homothetic (Fieler, 2011; Caron et al., 2014; Fajgelbaum and Khandelwal, 2016; Adao et al., 2017).

1 The Hamilton method

We assume the budget share of household h spends on food is:[3]

$$\frac{e_y^{hf}}{m_y^h} = \alpha + \gamma \ln \frac{P_y^f}{P_y^n} + \beta \ln \frac{m_y^h}{P_y} \tag{1}$$

Here m_y^h is total nominal expenditures, e_y^{hf} is nominal expenditures on food, P_f is the price index for food, and P_n is the price index for non-food. The price index P_y deflates nominal income. Deaton and Muellbauer (1980) show that this relationship is a first-order local approximation for any demand system, and show how one can calculate P_y given data on prices of goods categories,[4]

$$\ln P_y = \alpha_0 + \left(\alpha_f \ln P_y^f + \alpha_n \ln P_y^n\right) + \frac{1}{2}\gamma \left(\ln P_y^f - \ln P_y^n\right)^2 \tag{2}$$

The parameter β governs how the share of expenditures on food reacts to changes in log real income. Put another way, β is the slope of the Engel curve.

With an eye to empirics, we do not usually calculate the price level itself, but rather changes in the price level relative to a base year $P_y = P_0(1 + \pi_y)$:

$$\frac{e_y^{hf}}{m_y^h} = \alpha + \gamma \ln \frac{\left(1 + \pi_y^f\right) P_0^f}{\left(1 + \pi_y^n\right) P_0^n} + \beta \ln \frac{m_y^h}{P_0(1 + \pi_y)} \tag{3}$$

[3]For ease of exposition, we do not include household observables which affect expenditure share on food, such as number of children. We include such controls below in our estimation section.

[4]While we assume that P_y is the correct deflator of income for our expenditure function, it should not be interpreted as a cost of living index. That is, deflated income is not equal to utility. Indeed, as long as $\beta \neq 0$, our expenditure function implies that there is no scalar cost of living index, since households at different points in the income distribution buy different bundles of goods. For a thorough discussion of this point, and how one can compare a price index similar to ours to the cost of living, see Almås et al. (2018).

Suppose a statistical agency would like to calculate the inflation rate π_y. This is a complicated calculation, and they ultimate observe a biased measure of inflation $(1+\pi_y^b) = (1+\pi_y)b_y$. Replacing the true price index in (3) with the biased measure gives us:

$$\begin{aligned}\frac{e_y^{hf}}{m_y^h} &= \alpha + \gamma \ln \frac{\left(1+\pi_y^f\right) P_0^f}{\left(1+\pi_y^n\right) P_0^n} + \beta \ln \frac{m^h b_y}{P_0(1+\pi_y^b)} \\ &= \tilde{\alpha} + \gamma \ln \frac{\left(1+\pi_y^f\right)}{\left(1+\pi_y^n\right)} + \beta \ln \frac{m_y^h}{(1+\pi_y^b)} + \beta \ln b_y \end{aligned} \tag{4}$$

Intuitively, if households observed in different years but with the same level of real expenditures have different shares of spending on food, our inflation measure is likely biased.

In standard household budget survey data, we observe both expenditure shares and total nominal expenditures. Given a price indexes, and adding a noise term we assume is uncorrelated with regressors, we can use data on observed household expenditures to estimate the following relationship using a fixed effects OLS regression:

$$\frac{e_y^{hf}}{m_y^h} = \tilde{\alpha} + \gamma \ln \frac{\left(1+\pi_y^f\right)}{\left(1+\pi_y^n\right)} + \beta \ln \frac{m_y^h}{(1+\pi_y^b)} + \delta_y + \varepsilon_y^h \tag{5}$$

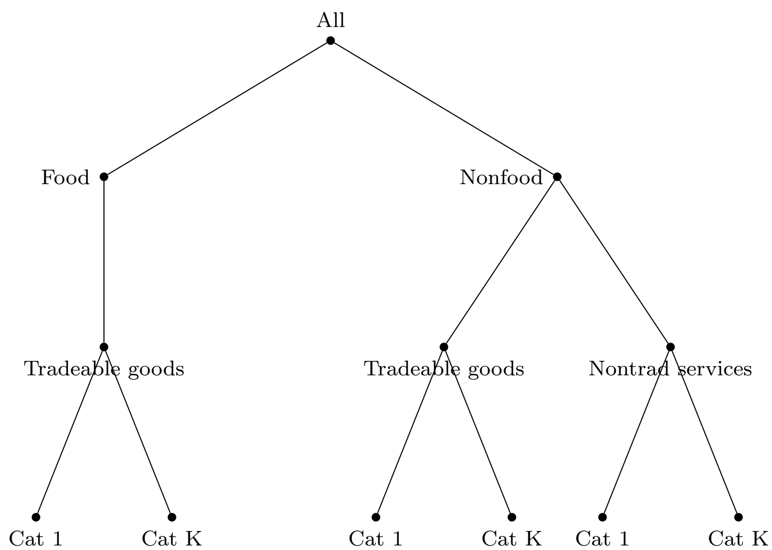
Our estimate $\hat{\beta}$ is a consistent estimator of the slope of the Engel curve for food, and the fixed effects $\hat{\delta}_y$ will be consistent estimators of annual inflation bias $\beta \ln b_y$ as the number of annual observations goes to infinity. Our primary empirical exercises will be based on Equation (5).

2 Constructing a Trade-Based Price Index

At its most abstract level, a price index is an weighted mean of prices on individual goods or varieties. In constructing these means, one challenge is that we have data on trade unit values but not on domestic prices except what we get from national statistical agencies. Our approach will be to construct price indexes common in the trade literature. Often the trade unit values will only allow us to construct import price indexes, which must be combined with prices on domestically produced tradeables and non-tradeables in order to create a consumer price index. In this section we develop a general model which puts enough structure on this problem for us to combine trade prices with domestic prices without directly observing domestic prices.

Since we focus on downward sloping food Engel curves, we build non-homothetic behavior into a two sectors in the upper tier of demand, Food and Non-food. For lower nests of preferences, we will focus on some popular price indexes in the trade literature that are derived from homothetic demand.

Figure 1: Structure of goods and services demand



In each sector s there are multiple categories of goods $k \in \Omega_s$. The sector-level price index in country i is defined as:

$$P_{it,s} = P_s\Big(\{P_{it,sk}/\phi_{it,sk}\}_{k\in\Omega_s}\Big) \tag{6}$$

where $P_{it,sk}$ is the price of category k in sector s country i time period t with an associated demand shifter $\phi_{it,sk}$. Here, our fairly general assumptions are that the set of categories Ω_s is location- and time-invariant, function $P_s(.)$ is invertible, demand is homothetic across categories within a sector, and log demand shifters move log prices additively. The price index of category k is in turn defined over varieties within that category,

$$P_{it,sk} = P_{sk}\Big(\{p_{it,sk\omega}/\phi_{it,sk\omega}\}_{\omega\in\Omega_{it,sk}}\Big) \tag{7}$$

Here, ω is a variety within sector-category sk, and $\Omega_{it,sk}$ is the set of available varieties to households in country i period t reflecting products entry and exit. We define varieties based on available data and specifications of our interest. For example, varieties might be differentiated by origin country, or by exporter firms, or by importing firm, or a combination of these attributes.

We assume that substitution patterns are such that the uppermost sectors are Food (f) and Non-food. Non-food is partitioned into tradeable and non-tradeable sectors (n) and (s). We illustrate the nesting structure in Figure 1.

Sector-level Prices. A large body of literature has examined two prominent features of consumer behavior with respect to tradeables (goods) and nontradeables (services). This literature asks what extent goods and services are complements, and the extent to which services are more income elastic than goods.[5] To incorporate these two margins of consumer behavior, we consider a non-homothetic CES that nests both standard CES and Cobb-Douglas.[6] If we know the price of tradeables with sector s, as well as their expenditure share, we don't need the prices of non-tradeables within a sector. The price index of sector s is:

$$P_{it,s} = (\pi^T_{it,s})^{\frac{1}{\gamma_s - \varepsilon^T_s}} (P^T_{it,s})^{\frac{\gamma_s - 1}{\gamma_s - \varepsilon^T_s}} \tag{9}$$

where ε^T_s and ε^N_s are income elasticities, and γ_s is the elasticity of substitution between the tradeable goods and nontradeable services within sector s. The intuition for this equation is that, conditional on tradeable prices, if the spending share on tradeables is higher, the price of non-tradeables must also higher.

Across Varieties within Categories. Next we go one nest down to calculate the price of tradeables within a sector from observed prices of varieties. A particular case of equation (7) that is widely used across the trade literature is a demand system based on constant elasticity of substitution (CES),

$$P_{it,sk} \quad = \quad \left[\sum_{\omega \in \Omega_{it,sk}} \Big(p_{it,sk\omega} / \phi_{it,sk\omega} \Big)^{1-\sigma_{sk}} \right]^{1/(1-\sigma_{sk})} \tag{10}$$

where $\sigma_{sk} > 0$ is the elasticity of substitution between varieties within a category.

To start our analysis, we add some more structure to the way that categories within sectors are aggregated, in order to connect the theory to data on international trade. Specifically, we assume all varieties within sector-category sk from country j (purchased by households in country i time period t) have the same unit value. Furthermore, assuming multiplicative trade costs, let $\tau_{ijt,sk}$ augment both trade costs and demand shifters for transactions associated to (ijt, sk), and $p_{jt,sk}$ be the unit value at the original production location j. We normalize $\tau_{iit,sk} = 1$, meaning that $\tau_{jit,sk}$ is interpreted as trade costs and exogenous demand for exporting country j relative to those of domestic. Putting

[5] See Herrendorf, Rogerson and Valentinyi (AER, 2013), Cravino and Sotelo (AEJ Macro, 2019), Comin, Lashkari, Mestieri (2019), and many papers they cite back to Baumol (AER, 1967)

[6] Specifically, CES demand implies the following price index for sector s,

$$P_{it,s} \quad = \quad \left[\Big(\frac{E_{it,s}}{P_{it,s}}\Big)^{\varepsilon^T_s - 1} \Big(\frac{P^T_{it,s}}{\phi^T_{it,s}}\Big)^{1-\gamma_s} + \Big(\frac{E_{it,s}}{P_{it,s}}\Big)^{\varepsilon^N_s - 1} \Big(\frac{P^N_{it,s}}{\phi^N_{it,s}}\Big)^{1-\gamma_s} \right]^{\frac{1}{1-\gamma_s}} \tag{8}$$

Here, $E_{it,s}$ is expenditure on sector s, $P^T_{it,s}$ and $P^N_{it,s}$ are prices of tradeables and nontradeables, ε^T_s and ε^N_s are income elasticities, and γ_s is the elasticity of substitution between the tradeable goods and nontradeable services within sector s. If $\varepsilon^N_s > \varepsilon^T_s$ then the nontradeable service is a luxury and tradeable good a necessity, and if the two equal one the system collapses to standard CES. In addition, if $\gamma_s < 1$, the tradeable and nontradeable sub-sectors are complement, if $\gamma_s > 1$, they are substitutable, and if $\gamma_s = 1$ and $\varepsilon^N_s = \varepsilon^T_s = 1$, the system collapses to Cobb-Douglas.

these together, the share of expenditure on country j and the price index implied by equation (10) are given by

$$\pi_{ijt,sk} = \frac{M_{ijt,sk}(\tau_{ijt,sk}p_{jt,sk})^{1-\sigma_{sk}}}{P_{it,sk}^{1-\sigma_{sk}}} \tag{11}$$

$$P_{it,sk} = \left[\sum_{j\in J} M_{ijt,sk}\Big(\tau_{ijt,sk}p_{jt,sk}\Big)^{1-\sigma_{sk}}\right]^{1/(1-\sigma_{sk})} \tag{12}$$

Here, $M_{ijt,sk}$ is the number of varieties in sector-category sk that are available to country i in time period t from supplying country j. Another interpretation is that $M_{ijt,sk}$ reflects domestic factor productivity together with $p_{it,sk}$ as the unit cost of production.[7]

The domestic price index $P^D_{it,sk}$ and the import price index $P^M_{it,sk}$ are given by:

$$P^D_{it,sk} = M_{iit,sk}^{1/(1-\sigma_{sk})}p_{it,sk} \tag{13}$$

$$P^M_{it,sk} = \left[\sum_{j\in J, j\neq i} M_{ijt,sk}\Big(\tau_{ijt,sk}p_{jt,sk}\Big)^{1-\sigma_{sk}}\right]^{1/(1-\sigma_{sk})} \tag{14}$$

Let $\pi^D_{it,sk}$ and $\pi^M_{it,sk}$ denote domestic and import expenditure shares. The price index can be then written as:

$$P_{it,sk} = \left[(P^D_{it,sk})^{1-\sigma_{sk}} + (P^M_{it,sk})^{1-\sigma_{sk}}\right]^{1/(1-\sigma_{sk})}$$

Next we connect the overall sector-category price index $P_{it,sk}$ to just the import price index $P^M_{it,sk}$ and the expenditure share on imports. This will allow us to use only firm- or product-level unit values that are abundant in trade data to calculate the overall price index. In the following, for a generic variable x_t, let $\widehat{x}_t = x_t/x_{t-1}$. Let the expenditure share on imports be denoted by $\pi^M_{it,sk} \equiv \sum_{j\neq i}\pi_{ijt,sk}$. Then the overall price index is a combination of the expenditure share and the import price index,

$$P_{it,sk} = \left(\pi^M_{it,sk}\right)^{\frac{1}{\sigma_{sk}-1}} P^M_{it,sk}$$

The import price index is given by

$$\widehat{P}^M_{it,sk} = \left(\widehat{\lambda}_{it,sk}\right)^{1/(1-\sigma_{sk})} \prod_j \left(\widehat{p}_{ijt,sk}\right)^{b_{ijt,sk}} \tag{15}$$

[7]In basic Armington and Eaton-Kortum models, M reflects exogenous demand for quality or the exogenous state of technology. If we allow entry and exit, more in line with Krugman or Melitz, then M is endogenous. For the sake of our analysis, it does not matter if M is exogenous or endogenous because we do note model the pricing rule and the entry/exit decision for firms, and because even if we follow Eaton-Kortum, technology may still change over time in ways that show up in data but are not modeled in their static theory.

Here $\lambda_{it,sk}$ is the share of spending on varieties which are imported in both t and $t-1$, which are the only varieties for which we have price changes $\widehat{p}_{ijt,sk}$. $b_{ijt,sk}$ are the Sato-Vartia weights,

$$b_{ijt,sk} = \frac{(x_{ijt,sk} - x_{ijt-1,sk})/(\ln x_{ijt,sk} - \ln x_{ijt-1,sk})}{\sum_{l \in l \neq i}(x_{ilt,sk} - x_{ilt-1,sk})/(\ln x_{ilt,sk} - \ln x_{ilt-1,sk})} \tag{16}$$

where $x_{ijt,sk}$ is expenditures on variety j as a share of all expenditures on varieties available at both t and $t-1$. Connecting these two equations,

$$\widehat{P}_{it,sk} = \left(\widehat{\pi}^{M}_{it,sk}\right)^{\frac{1}{\sigma_{sk}-1}} \left(\widehat{\lambda}_{it,sk}\right)^{-\frac{1}{\sigma_{sk}-1}} \prod_j \left(\widehat{p}_{ijt,sk}\right)^{b_{ijt,sk}} \tag{17}$$

If we have trade data, at a disaggregated level, we can construct $\widehat{\lambda}$ and $\widehat{p}$. If we have data on domestic production at the sector level, we can construct $\widehat{\pi}$.

3 Data

We have several primary data sources. To construct price indexes, we use trade data are taken from the United Nationals Comtrade database, and production data from OECD STAN. These data sources are both free and easily accessible online. In order to evaluate the quality of price indexes using the Data on Danish household expenditures are data from the Danish Household Budget Survey, and data on American household expenditures are taken from the Panel Survey of Income Dynamics. These data sets have been documented extensively elsewhere, but even so we briefly describe them in turn.

We use UN Comtrade data to construct bilateral trade flows by product, which we need to construct our trade-based price indexes. Comtrade data is compiled by the United Nations Statistics Division, and is based on commodity-partner level trade data reported to the United Nations by 170 reporting countries. We use data from 1993 to 2012, which is easily available for download through the API available on the UN's website. Several transformations of the data are performed by the United Nations. All values are converted to current US dollars using exchange rates from the reporting countries, and commodity codes are converted to be consistent over the entire data period.

In order to construct import shares by sector/category, we use OECD STAN data. This data is available for download on the OECD's data dissemination website OECD.Stat. The coverage is 1990-2016, although it can vary by country and industry. We have all countries and industries from 2000-2016. The data include both quantity and value information for domestic production by ISIC industry. Data ultimately come from national statistics offices, but the OECD also imputes missing data.

In order to calculate Danish household expenditures and construct demographic controls, we use the Danish Household Budget Survey. We have access to microdata from 1996-2016. This data set is used by Denmark Statistics to construct expenditure shares for price index calculation. Around 1100 households are surveyed annually. The data are collected according to Eurostat guidelines, and in particular the expenditure categories are reported at the COICOP5 level. Households are uniformly randomly selected to take part in the survey, and then weights are calculated afterwards to make the sample representative of the entire Danish population.

Finally, American household expenditures are taken from the Panel Survey of Income Dynamics. Unlike the Danish data, for this data we observe the same households across many years, although we do not use this dimension. The data is freely available from the University of Michigan's Institute for Social Research. The number of households changes over time as the survey tracks descendants of the original 4800 families surveyed in 1968. The coverage goes from 1968 to 2017. The data contains information on household characteristics, nominal income, and expenditures on broad consumption categories.

3.1 Descriptive Statistics

Since the UN Comtrade data and OECD STAN data are standard sources for international trade and production data at the commodity level, we do not calculate descriptive statistics for them here. We rather focus on the less well-known Danish Household Budget Survey.

3.1.1 Danish Household Budget Survey

Tables 1-4 contain descriptive statistics from our sample. Denmark Statistics defines the household head to be the female if a household is composed of an opposite gendered pair. Otherwise, the household head is the oldest member of the household. This definition is why Table 2 reports that 80% of household heads are female, and nearly all spouses are male. The observation numbers in Table 4 are lower than in the other tables because some people have no reported education. The employment numbers in Table 2 are relatively low, only 58% of heads and 72% of spouses. Of those less than 40% are employed full-time. This is because many households are older – the average age of both heads and spouses is 50 – and in this period early retirement was common (foertidspension). One might be surprised that Table 3 reports the highest fraction of our sample comes from East Jutland rather than Copenhagen. East Jutland is the location of Denmark's second largest city, Aarhus. Even so, Copenhagen is by far the largest Danish city. The reason why it does not look that way in the table is because the commuting zone includes parts of North and East Sealand, which are counted separately in the table.

More directly related to our project, Table 1 reports that on average Danish household expenditures were around half of Danish gross income. This might seem low, but recall that Danish taxes are among the highest in the world. On average labor income is taxed at around 40%, so actually Danes are spending a substantial portion of their net income. Annual expenditure on food in our sample was around 37,000 Danish Kroner. This is around 15% of total expen-

	Denmark		United States	
	Mean	SD	Mean	SD
Persons in household	2.12	1.23		
Number of children	1.64	0.64	0.97	1.16
Gross Income	512,368	404,690	52,258	43,583
Expenditure	243,896	140,056		
Net Real Income			11,969	8132
Nominal expenditure on Food	37,536	25,041	4695	2808
Expenditure Share on Food	0.159	0.068	0.124	0.081
Obs	16,328		62,406	

Table 1: Household Statistics

	Denmark				United States			
	Head		Spouse		Head		Spouse	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	0.80	0.40	0.00	0.05	0	0	0	0
Age	49.58	18.08	50.47	15.54	47.05	15.25	44.60	14.74
Fraction employed	0.58	0.49	0.72	0.45	0.80	0.40	.63	.48
Of employed, over 35 hours	0.38	0.48	0.34	0.47	0.68	0.47	0.34	0.47
Of employed, work hours	33.98	9.98	38.66	9.80	46.14	17.57	34.40	17.56
Obs	16,340		16,340		62,406		62,406	

Table 2: Head and Spouse Statistics

ditures. If we take the mean of household individual expenditure shares, we get a similar number, around 16%.

3.1.2 PSID

Tables 1,2, and 5 contain descriptive statistics for our PSID sample. American household heads are all males by definition. Our sample of American households are a little younger than Danish households, and our American households have fewer children. The mean nominal household income is 52,258 USD. Food expenditures as a share of total expenditures is 0.124. This is lower than in Denmark, potentially implying that Americans are wealthier on average. On the other hand, it could also be because our sample of Danes has more children. Consistent with stereotypes, Americans tend to work more, and conditionally on working are more likely to be full time and also to work more hours.[8] While education distributions are difficult to compare with the different educational systems, the numbers look roughly similar to Denmark, with 51% of heads and 55% of spouses having a high school education or below.

[8]We only had annual hours for American workers, so to get a weekly hour number we divided by 50. If anything, this should be a lower bound on weekly hours, since some American workers get more than two weeks of vacation.

Location	Percent
East Jutland	28.1
Copenhagen	24.6
North Jutland	11.0
West/South Sealand	9.8
Fyn	7.2
North Sealand	6.8
East Sealand	3.8
Bornholm	1.5
Obs	15,958

Table 3: Danish Household Sample Location

Education	Head	Spouse
Primary school	41.71	39.12
Vocational edu	14.20	14.28
Business secondary edu	0.66	0.44
General secondary edu	9.18	6.83
Short higher edu	4.67	6.73
Medium-long higher edu	20.82	17.49
Bachelor edu	2.28	1.90
Long higher edu	6.00	12.07
Research training	0.48	1.13
Obs	11,392	5,350

Table 4: Danish Household Highest Completed Education

	Head	Spouse
Less than High School	18.84	15.64
High School	32.05	39.85
Some College	19.82	20.80
College	17.40	14.95
Graduate Education	11.90	8.76
N	62406	62406

Table 5: US Household Highest Completed Education

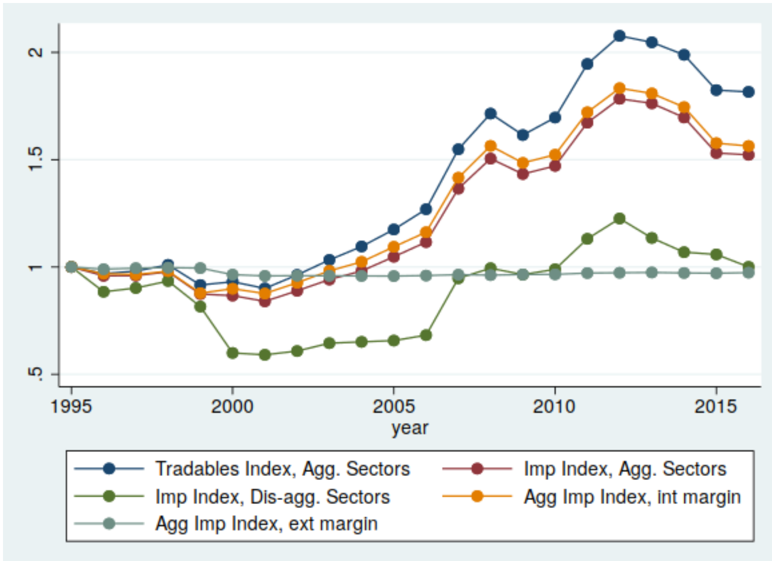


Figure 2: Disaggregated and Aggregated Import Indexes

4 Empirical Results

4.1 Empirical Price Index Construction

In Section 2 we described how to construct a price index based on trade data. Our trade data come at a very disaggregated level. If we partition our trade data into 255 SITC-3 categories, we can construct the import price index (15) using elasticities of substitution σ_{sk} estimated and reported by Broda and Weinstein (2006). In all of our trade-based price indexes, varieties will be imports at the 8-digit HS code, differentiated by country of origin. We will sometimes refer to the import price index calculated this way as the dissaggregated import price index. To proxy domestic prices using equation (17), however, we need production data by year and category in order to calculate category-level import shares $\pi^M_{it,sk}$. Unfortunately this data is not available with such a fine level of disaggregation. We therefore also partition our trade data into 11 categories available in our production data. For these categories, we can construct import shares by year, but we do not have an estimate of σ_{sk}. For now we assume that $\sigma_{sk} = 4$ for all categories. We will refer to the import price index calculated with our 11 categories as the aggregated import price index, and refer to the import-share adjusted index calculated using (17) as the tradeables price index. We combine all import indexes and the tradeables price index with a BLS provided services price index using BLS CPI weights on goods and services.

Figure 2 compares the trade-based indexes described above. The disaggregated import index falls into the early 2000's and is more or less flat after 2007. The aggregated import index is relatively flat until 2007, and then rises substantially, rising by 70% in the early 2010's, and then falling a bit again. The flat line close to zero is the contribution of the extensive margin to the aggregate index. Unsurprisingly, when the data are aggregated to this degree, the number of varieties does not vary much across years. Our tradeables price index, which is the aggregated import index augmented by changes in the import share, is higher yet again. This implies that domestic prices were increasing even more quickly than import prices in this period.

4.2 US Results

In this section we evaluate the price index described in Section 2 using the Hamilton regression described in Section 1. We use a slightly modified version of equation (1) to allow for controls which surely matter for food expenditures shares, such as number of children. We also add an error term. This modified version is:

$$\frac{e_y^{hf}}{m_y^h} = \tilde{\alpha} + \gamma \ln \frac{\left(1+\pi_y^f\right)}{\left(1+\pi_y^n\right)} + \beta \ln \frac{m_y^h}{\left(1+\pi_y^b\right)} + \lambda X_y^h + \delta_y + \varepsilon_y^h \tag{18}$$

Our results for the American PSID data are contained in Tables 6 and 7. Table 6 contains log income and the other demographic regressors. The estimates in this table do not depend on how we deflate log income, so they are the same in all of our specifications. The coefficient on log income is the slope of the Engel curve, which we estimate to be -0.0768. This is similar to the rule of thumb food Engel curve slope of -0.1. Households with older heads, and heads that work more tend to spend a higher share of expenditures on food. Children reliably increase the share of household expenditures on food. Nothing else is statistically significant.

Table 6 is the main table of interest for our US estimation. It contains all variables which change depending on specification. The first column corresponds to the model deflating income by officially provided CPI. The second column deflates income with the disaggregated import index. The third deflates using the aggregated import index. The fourth deflates income with our overall trade-based consumer price index. Finally the last row is simply nominal income. The first row is the log relative price of food and non-food, which we expect to have a negative effect on food expenditure share. This prediction is upheld in all specifications except the undeflated nominal income column.

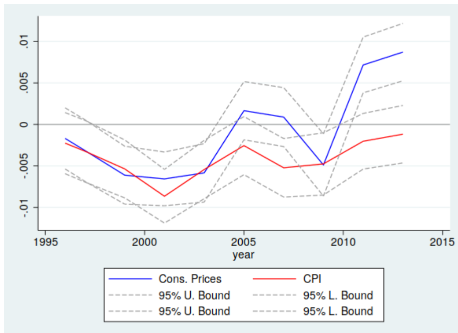
To ease in interpretation, we plot results on year dummies from Table 6 in Figure 3. Recall that we expect these shifters to be zero if we have correctly deflated nominal expenditures. Subfigure 3a contains constant prices, where we set $P_y = 1$ for all years, and also the CPI specification. The year coefficients in our regression are first negative, and then positively significantly different from zero for constant prices. If people's nominal expenditures were real, in order to match our data on food expenditure shares, we would have to shift the Engel curve up as years went by. That is, conditional on expenditures people would be spending a larger share of their expenditures on food as years passed. If we

	Food expenditure share
ln inc.	-0.0768***
	(0.00124)o
age spouse	-0.000152
	(0.0000962)
age head	0.000374***
	(0.0000949)
age spouse	-0.000152
	(0.0000962)
childrn	0.0148***
	(0.000439)
hrs head	0.00135***
	(0.000359)
hrs wife	-0.00335
	(0.000322)
edu head	-0.0000574
	(0.000183)
edu wife	0.0000411
	(0.000146)
cons	1.187***
	(75.50)
N	24043
R-sq	0.393
Year FE	Yes

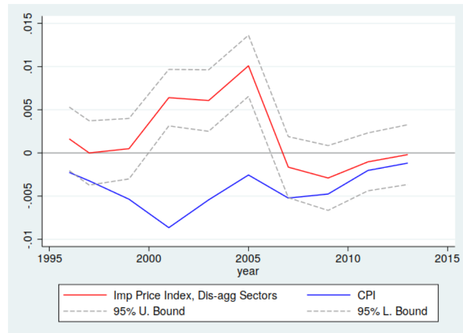
Table 6: Regression results: Common to all price indexes

	Food expenditure share				
	CPI	Dissagg. Import P	Agg. Import P	Trade-based P	Nominal
rel P food	-0.0169**	-0.00102	-0.0273***	-0.0398***	0.0912***
	(0.00617)	(0.00617)	(0.00617)	(0.00618)	(0.00645)
1996	-0.00226	0.00163	-0.000216	-0.000320	-0.00169
	(0.00188)	(0.00188)	(0.00188)	(0.00188)	(0.00188)
1997	-0.00323	-0.000000965	-0.000819	-0.000984	-0.00315
	(0.00189)	(0.00190)	(0.00190)	(0.00190)	(0.00189)
1999	-0.00536**	0.000493	0.000424	-0.00000904	-0.00612***
	(0.00178)	(0.00179)	(0.00179)	(0.00179)	(0.00178)
2001	-0.00865***	0.00640***	-0.000652	-0.00140	-0.00656***
	(0.00165)	(0.00167)	(0.00166)	(0.00165)	(0.00165)
2003	-0.00543**	0.00607***	0.0000237	-0.000601	-0.00584**
	(0.00179)	(0.00182)	(0.00180)	(0.00180)	(0.00179)
2005	-0.00255	0.0101***	0.00138	0.0000805	0.00165
	(0.00179)	(0.00181)	(0.00179)	(0.00179)	(0.00179)
2007	-0.00523**	-0.00165	-0.00712***	-0.00846***	0.000883
	(0.00180)	(0.00181)	(0.00180)	(0.00179)	(0.00181)
2009	-0.00476*	-0.00290	-0.00654***	-0.00650***	-0.00488*
	(0.00191)	(0.00191)	(0.00191)	(0.00191)	(0.00191)
2011	-0.00203	-0.00103	-0.00606***	-0.00756***	0.00716***
	(0.00171)	(0.00171)	(0.00171)	(0.00171)	(0.00171)
2013	-0.00118	-0.000195	-0.00556**	-0.00661***	0.00871***
	(0.00177)	(0.00177)	(0.00177)	(0.00177)	(0.00176)

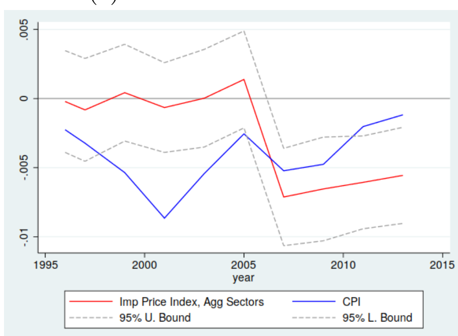
Table 7: Regression results: Year fixed effects



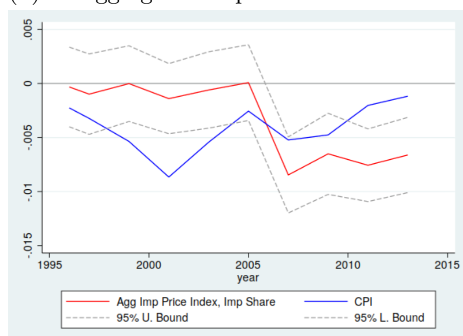
(a) No inflation vs US CPI



(b) Disaggregated Import Index vs US CPI



(c) Aggregated Import Index vs US CPI



(d) Tradeables Index vs US CPI

Figure 3: Year Dummies

instead assume that the Engel curve is fixed, the pattern in Figure 2 implies that constant prices first understated, and then overstated real expenditures. The year dummies in the CPI specification are also significantly different from zero, but remain negative for the whole period. Our estimates imply that we must shift the Engel curve down to match the data, implying that people spent a smaller share of their expenditures on food, conditional on income, than they did in 1995, our base year. Assuming that the Engel curve is fixed, these estimates imply that CPI is too high, understanding people's true real expenditures.

Subfigure 3b is our disaggregated import index. Since this is purely an import index, it is somewhat surprising that it overstates real income in the early 2000's, a period of trade expansion. Our aggregated import index does a better job in the early 2000's, but does less well after 2007, understating real income even more than BLS-reported CPI. Adding the import shares to make the aggregate index a measure of overall CPI does not change this conclusion. The disaggregated import index and the aggregated import index do better in exactly opposite time periods, a result which we intend to explore further. The takeaway from this exercise is that even though imports make up less than 20% of American GDP, our calculated import price indexes are more accurate than the standard government calculated CPI. That is, import price indexes better predict food expenditure shares given the estimated decreasing Engels curve in food.

4.3 Danish Results

TBA

5 Conclusion

TBA

References

Adao, R., Costinot, A., and Donaldson, D. (2017). Nonparametric counterfactual predictions in neoclassical models of international trade. *American Economic Review*, 107(3):633–89.

Almås, I. (2012). International income inequality: Measuring ppp bias by estimating engel curves for food. *American Economic Review*, 102(2):1093–1117.

Almås, I., Beatty, T. K., and Crossley, T. F. (2018). Lost in translation: What do engel curves tell us about the cost of living?

Arkolakis, C., Costinot, A., and Rodriguez-Clare, A. (2012). New trade models, same old gains? *American Economic Review*, 102(1):94–130.

Bai, L. and Stumpner, S. (2019). Estimating us consumer gains from chinese imports. *American Economic Review: Insights*, 1(2):209–24.

Banks, J., Blundell, R., and Lewbel, A. (1997). Quadratic engel curves and consumer demand. *Review of Economics and statistics*, 79(4):527–539.

Broda, C. and Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly journal of economics*, 121(2):541–585.

Caron, J., Fally, T., and Markusen, J. R. (2014). International trade puzzles: A solution linking production and preferences. *The Quarterly Journal of Economics*, 129(3):1501–1552.

Chai, A. and Moneta, A. (2010). Retrospectives: engel curves. *Journal of Economic Perspectives*, 24(1):225–40.

Deaton, A. and Muellbauer, J. (1980). An almost ideal demand system. *The American economic review*, 70(3):312–326.

Deaton, A. S. and Muellbauer, J. (1986). On measuring child costs: With applications to poor countries. *Journal of Political Economy*, 94(4):720–744.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Fajgelbaum, P. D. and Khandelwal, A. K. (2016). Measuring the unequal gains from trade. *The Quarterly Journal of Economics*, 131(3):1113–1180.

Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *The American Economic Review*, pages 157–177.

Fieler, A. C. (2011). Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation. *Econometrica*, 79(4):1069–1101.

Hamilton, B. W. (2001). Using engel's law to estimate cpi bias. *American Economic Review*, 91(3):619–630.

Nakamura, E., Steinsson, J., and Liu, M. (2016). Are chinese growth and inflation too smooth? evidence from engel curves. *American Economic Journal: Macroeconomics*, 8(3):113–44.

Redding, S. and Weinstein, D. E. (2018). Accounting for trade patterns. *Princeton University, mimeograph.*