NBER WORKING PAPER SERIES

A UNIFIED APPROACH TO ESTIMATING DEMAND AND WELFARE

Stephen J. Redding
David E. Weinstein

Working Paper 22479
http://www.nber.org/papers/w22479

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2016, Revised March 2017

We particularly thank Dave Donaldson, Pablo Fajgelbaum, Rob Feenstra, Pete Klenow, and Nick Bloom for helpful comments. We are also grateful to many other colleagues and conference and seminar participants for their helpful comments. We also thank Mark Greenan, Mathis Maehlum and Dyanne Vaught for excellent research assistance. All results are calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business. Responsibility for the results, opinions, and any errors lies with the authors alone. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Stephen J. Redding and David E. Weinstein. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

A Unified Approach to Estimating Demand and Welfare
Stephen J. Redding and David E. Weinstein
NBER Working Paper No. 22479
August 2016, Revised March 2017
JEL No. D11,D12,E01,E31

ABSTRACT

Measuring aggregate price and welfare movements is central to international and macro economics. Existing approaches assume constant preference parameters for each good, which rules out demand shocks for individual goods. However, micro data display substantial shifts in expenditure shares conditional on prices, highlighting the empirical relevance of demand shocks. We develop a unified approach that allows for these demand shocks for individual goods, while preserving a money-metric utility function that can be used for comparisons of welfare over time. We implement our approach for one of the most widely-used preference structures in economics: constant elasticity of substitution (CES) preferences. We derive a unified price index that is exact for these preferences, allows for product entry/exit, demand shocks for individual goods, and nests all major existing economic and statistical approaches. We show that abstracting from demand shocks introduces a "consumer-valuation bias," which is analogous to the well-known "substitution bias," and results in a substantial overestimate of the increase in the cost of living over time. We develop a new "reverse-weighting" estimator for the elasticity of substitution between goods, which we use to derive upper and lower bounds for the true elasticity that permit set identification regardless of the correlation between demand and supply shocks.

Stephen J. Redding
Department of Economics
and Woodrow Wilson School
Princeton University
Fisher Hall
Princeton, NJ 08544
and NBER
reddings@princeton.edu

David E. Weinstein
Columbia University, Department of Economics
420 W. 118th Street
MC 3308
New York, NY 10027
and NBER
dew35@columbia.edu

1 Introduction

Measuring price aggregates is central to international and macro economics, which depend crucially on being able to distinguish real and nominal income.[1] However, constructing measures of the aggregate cost of living raises a number of challenges, of which two are particularly problematic. First, all existing price aggregates assume that the preference parameters for each good are time invariant, which rules out the possibility of demand shocks arising from changes in consumer tastes. However, a large microeconometric literature in economics is predicated on the very assumption of such demand shocks. Indeed, explaining observed data on prices and expenditure shares within the framework of a model typically requires the assumption of non-zero demand shocks. Therefore, there is a disconnect between the assumption of no demand shocks in existing price indices and the properties of the micro data on prices and expenditure shares from which they are constructed. Second, measuring aggregate prices often requires knowing demand-side parameters for many industries (as in the variety-adjusted price index proposed by Feenstra (1994)). In estimating these parameters, researchers are typically concerned about the correlation between demand and supply shocks, but do not have access to the detailed information required to construct valid instrumental variables for every industry in the aggregate economy.

In this paper, we develop a new approach to aggregate welfare measurement that addresses both of these challenges. We refer to it as the unified approach, because it both exactly rationalizes the observed micro data on prices and expenditure shares as an equilibrium outcome, while also permitting exact aggregation and comparisons of national welfare over time, thereby unifying micro and macro. A key advantage of this approach relative to conventional price indices is that we incorporate shifts in demand for surviving goods and the entry and exit of goods over time. Both are central features of micro data. We show how to aggregate from these micro data to a money-metric price index, which depends solely on prices and expenditure shares. Our approach does not require an outside good and is easy to implement for a broad range of sectors when only data on prices and expenditure shares are available. Hence, our approach is well suited for macro and trade applications, in which researchers are concerned with the economy as a whole and wish to allow for general equilibrium effects across sectors. We show how functional form assumptions about preferences alone can be used to derive upper and lower bounds for demand-side parameters when valid instruments are not available. If demand and supply shocks are orthogonal, these bounds coincide, and we obtain point identification. If they are correlated, these bounds differ from one another and identify the set of parameter values consistent with the observed data and our assumptions about demand.

The central insight underlying our approach is to use the demand system to substitute out for the time-varying preference parameters for each good. As long as the demand system is invertible, these unobserved preference parameters can be uniquely determined from the observed data on prices and expenditure shares (up to a choice of units in which to measure these demand parameters). We thus obtain a money-metric expression for the change in the cost of living that depends solely on these observed prices and expenditure shares. We focus throughout most of our analysis on constant elasticity of substitution (CES) preferences,

[1] Recent contributions to the measurement of the cost of living and aggregate productivity across countries and over time include Bils and Klenow (2001), Hsieh and Klenow (2009), Jones and Klenow (2016), Feenstra (1994), Neary (2004) and Syverson (2016).

because there is little doubt that this is the preferred approach to modeling product variety in the macro, trade, and economic geography literatures. This preference structure also provides a tractable approach to incorporating both demand shocks for surviving goods and the entry and exit of goods. We show that the unified price index that we derive from these preferences nests all major economic and statistical approaches in the existing price index literature. These can all be viewed as special cases that impose particular parameter restrictions (on the elasticity of substitution), ignore particular moments of the data (e.g. entry and exit of goods), or neglect certain implications of the model (the demand system and unit expenditure function are derived from the same underlying preferences).

We show that abstracting from demand shocks for individual goods introduces a substantial bias into conventional price indices that we term the "consumer valuation bias." This bias is analogous to the well-known "substitution bias," according to which a Laspeyres price index tends to overstate the increase in the cost of living over time, because it does not take into account that consumers can substitute away from the goods whose prices have risen. In our framework, the demand parameters for individual goods enter inversely with prices (so that consumers care about demand-adjusted prices). Therefore, the consumer valuation bias reflects the fact that standard price indices assume that these demand parameters are constant, and hence do not take into account that consumers can substitute towards the goods that they like more. In our empirical application using U.S. bar-code data, we find that this consumer valuation bias is substantial. It is around the same magnitude as the bias in conventional price indices from abstracting from the entry and exit of goods and amounts to more than one percentage point per annum in our baseline specification.

To implement our approach for CES preferences, we develop a new "reverse-weighting" estimator of the elasticity of substitution between goods. We show that this estimator permits identification using only assumptions about demand and regardless of the correlation between demand and supply shocks. The intuition for identification comes from counting equations and unknowns. If we think about a dataset containing price and share changes for k goods, we have k unknowns (one unknown price elasticity and $k-1$ unknown values for each of the preference shocks given a normalization). However, we also have a system of k independent equations ($k-1$ independent demand equations and one equation for the change in the unit expenditure function). Therefore, the system is exactly identified. We show that this reverse-weighting estimator is consistent as demand shocks become small or as the number of goods becomes large and demand shocks are independently distributed. More generally, we show that this estimator can be used to derive upper and lower bounds for the true elasticity of substitution regardless of the correlation of demand and supply shocks. In both a Monte Carlo and our U.S. data, we find that these bounds are tight, and identify a narrow range of possible values for the elasticity of substitution consistent with the observed data and the assumption of CES preferences. Therefore, we find similar results for the magnitude of the consumer valuation bias, even when we consider our most conservative specification using the upper bound for the elasticity of substitution.

Although we largely focus on CES preferences because of their prevalence in economics, we also report two major extensions of our approach. First, we show that all our main results (including the consumer valuation bias) also hold for the flexible functional form of translog preferences. Second, we show that our approach encompasses versions of random utility models, such as Fréchet and logit, in which the aggregate behavior of consumers with idiosyncratic tastes is the same as that of a representative consumer with CES preferences. We use this isomorphism to allow for multiple types of consumers with different substitution elasticities and preference parameters for each good, which results in mixed logit specification.

Our paper is related to a number of strands of existing research. First, we build on a long line of existing research on price indexes. Price measurement in most national and international agencies is based on the "statistical approach" to price indexes developed by Dutot (1738), Carli (1764), and Jevons (1863). The methodologies developed in these papers form the foundation of 98 percent of all consumer price indexes generated by government statistical agencies (Stoevska 2008). Following Konüs (1924), economic theory has largely rejected the "statistical approach" to price measurement in favor of the "economic approach," which asserts that all price indexes should be derived from consumer theory and correspond to the unit expenditure function. The subsequent economic approach to price measurement, including Diewert (1976), Sato (1976), Vartia (1976), Lau (1979), Feenstra (1994), Moulton (1996), Balk (1999), Caves, Christensen and Diewert (1982), Neary (2004) and Feenstra and Reinsdorf (2007, 2010), has focused on exact and superlative index numbers that feature time-invariant demand parameters.

We show that our unified price index nests all major price indices in existing research. Thus, how economists and statistical agencies currently measure welfare can be understood in terms of an internally consistent approach that has been altered by ignoring data, moment conditions, and/or imposing particular parameter restrictions. For example, allowing the elasticity of substitution to differ from the Cobb-Douglas assumption of one produces the Sato-Vartia (1976) constant elasticity of substitution (CES) exact price index.[2] Introducing the entry and exit of goods over time generates the Feenstra-CES index (Feenstra (1994). Incorporating demand shocks for each good and estimating the elasticity of substitution using the assumption that these time-varying demand shifts for individual products cancel on average produces the unified index. Other paths are shorter. The Jevons (1863) index—a geometric average of price widely used as an input into many price indexes—is a special case of the unified price index when the elasticity of substitution is infinite. The unified index exactly corresponds to expected utility if consumers have heterogeneous random utility with extreme value distributions (e.g., Logit or Fréchet). Similarly, the Dutot (1738), Carli (1764), Laspeyres (1871) and Paasche (1875) indexes all can be derived from the unified approach by making the appropriate parameter restrictions. Finally, relaxing assumptions necessary to yield the Fisher (1922) and Törnqvist (1936) indexes, yields the broader class of quadratic mean price indexes. The Sato-Vartia index arises naturally in this class, and as we just discussed, yields the unified price index if it is generalized. In other words, many seemingly fundamentally different approaches to welfare measurement—e.g., Laspeyres and Cobb-Douglas indexes—are actually linked together via the unified approach.

Our study is also related to a more recent, voluminous literature in macroeconomics, trade and economic geography that has used CES preferences. This literature includes, among many others, Anderson and van Wincoop (2003), Antràs (2003), Arkolakis, Costinot and Rodriguez-Clare (2012), Armington (1969), Bernard, Redding and Schott (2007, 2011), Blanchard and Kiyotaki (1987), Broda and Weinstein (2006, 2010),

[2]The "Cobb-Douglas" functional form was first used by Wicksell (1898) and the price index was discovered by Konyus (Konüs) and Byushgens (1926). Cobb and Douglas (1928) applied it to U.S. data. For a review of the origins of index numbers, see Chance (1966).

Dixit and Stiglitz (1977), Eaton and Kortum (2002), Feenstra (1994), Helpman, Melitz and Yeaple (2004), Hsieh and Klenow (2009), Krugman (1980, 1991), Krugman and Venables (1995) and Melitz (2003). Increasingly, researchers in international trade and development are turning to bar-code data in order to measure the impact of globalization on welfare. Prominent examples of this include Handbury (2013), Atkin and Donaldson (2015), and Atkin, Faber, and Gonzalez-Navarro (2015), and Fally and Faber (2016). Our contribution relative to this literature is to derive an exact price index that allows for changes in variety and demand for each good, while preserving the property of a money-metric utility function.

Our work is also related to research in macroeconomics aimed at measuring the cost of living, real output, and quality change. Shapiro and Wilcox (1996) sought to back out the elasticity of substitution in the CES index by equating it to a superlative index. Whereas that superlative index number assumed time-invariant demand for each good, we explicitly allow for time-varying demand for each good, and derive the appropriate index number in such a case. Bils and Klenow (2001) quantify quality growth in U.S. prices. We show how to incorporate changes in quality (or subjective taste) for each good into a unified framework for computing changes in the aggregate cost of living over time and estimating the elasticity of substitution.

The remainder of the paper is structured as follows. Section 2 outlines our general approach without specifying a particular functional form for utility. Section 3 assumes CES preferences and derives our unified price index. Section 4 examines the relationships between this unified price index and the standard price indexes used by economists and statistical agencies. Section 5 incorporates heterogeneous groups of consumers with different substitution parameters. Section 6 shows how our unified approach can be used to estimate the elasticity of substitution, characterizes the asymptotic properties of this estimator, and reports Monte Carlo results for its finite sample performance. Section 7 uses detailed bar-code data for the U.S. consumer goods sector to illustrate our approach and demonstrate its quantitative relevance for measuring changes in the aggregate cost of living. Section 8 concludes.

2 The Unified Approach

We begin by outlining our general approach without specifying a particular functional form for utility. We assume that the expenditure function is homothetic and twice continuously differentiable and that the demand system is invertible. In particular, we consider the following homothetic unit expenditure function defined over a constant set of goods $k \in \Omega$ with $N = |\Omega|$ elements:

$$\mathbb{P}_t = \mathbb{P}\left(\mathbf{P_t}, \boldsymbol{\varphi}_\mathbf{t}, \boldsymbol{\sigma}\right), \tag{1}$$

where $\mathbb{P}_t$ is a scalar corresponding to the cost of obtaining a unit of utility; we denote vectors in bold such that $\mathbf{P_t}$ is the vector of prices of the N goods; $\boldsymbol{\varphi}_\mathbf{t}$ is the vector of preference or "demand" parameters for each good; and $\boldsymbol{\sigma}$ is the vector of parameters that control substitution between goods. We define the demand parameters for each good ($\boldsymbol{\varphi}_\mathbf{t}$) such that they enter the unit expenditure function in an inverse way to prices (so that the consumer cares about demand-adjusted prices, P_{kt}/φ_{kt}). Applying Shephard's lemma to the unit expenditure function (1), we obtain the system of equations for the expenditure shares for each good:

$$\mathbf{S_t} = \mathbb{S}\left(\mathbf{P_t}, \boldsymbol{\varphi}_\mathbf{t}, \boldsymbol{\sigma}\right), \tag{2}$$

where $\mathbf{S_t}$ is the vector of expenditure shares S_{kt} for each good k at time t.

We interpret the demand parameters (φ_{kt}) as capturing consumer tastes, so that they appear in both the unit expenditure function (1) and the demand system (2). In our empirical analysis, these parameters act as structural residuals that ensure that the model's predictions for expenditure shares in equation (2) are exactly equal to the observed values in the data. In principle, these demand parameters could also capture product quality (in which case they would again appear in both the equations for preferences and demand) or measurement error in expenditure shares (in which case they would only appear in the demand system). In our empirical application, we use bar-code data, which alleviates concerns about measurement error. Furthermore, it is rare for firms to use the same bar code for products with different observable characteristics, because of the problems that this would create for stock and inventory control. Therefore, changes in observable product characteristics result in the introduction of a new bar code rather than changes in product quality within bar codes. Hence we interpret shifts in expenditure shares conditional on prices as shifts in consumer tastes. We show later that our approach naturally accommodates the entry and exit of goods over time and is robust to mean-zero, log-additive measurement error in prices and/or expenditure shares.

Our assumption that the demand system is invertible corresponds to the assumption that there exists a one-to-one mapping from the model's substitution parameters (σ) and the observed data on prices and expenditure shares ($\mathbf{S_t}$, $\mathbf{P_t}$) to the unobserved demand parameters ($\boldsymbol{\varphi}_\mathbf{t}$) for each good k at time t up to a normalization (the scalar $\bar{\varphi}$) that corresponds to a choice of units in which to measure the demand parameters:

$$\boldsymbol{\varphi}_\mathbf{t} = \mathbb{S}^{-1}\left(\mathbf{S_t}, \mathbf{P_t}, \sigma, \bar{\varphi}\right). \tag{3}$$

As expenditure shares in equation (2) are homogeneous of degree zero in prices and demand, these demand parameters ($\boldsymbol{\varphi}_t$) can only be recovered from the data on prices and expenditure shares up to this normalization or choice of units (the scalar $\bar{\varphi}$).

We now use this general formulation to contrast our unified approach with the standard price index approach to measuring changes in the cost of living over time. The standard approach assumes that the demand parameters for each good are time-invariant ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega$) and constructs a price index that depends solely on observed prices and expenditure shares:

$$\Phi^P_{t-1,t} = \frac{\mathbb{P}_t}{\mathbb{P}_{t-1}} = \Phi^P\left(\mathbf{P_t}, \mathbf{P_{t-1}}, \mathbf{S_t}, \mathbf{S_{t-1}}\right), \tag{4}$$

where the superscript P is a mnemonic for price index. This price index (4) is money metric, in the sense that it depends solely on prices and expenditure shares, which are directly comparable over time. Examples of this standard approach include the Sato-Vartia price index, which is exact for CES preferences, and the Törnqvist price index, which is exact for translog preferences.

A major challenge for this existing price index literature is that the assumption of time-invariant preference parameters for each good ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega$) implies that demand curves never shift in the expenditure share system (2). However, the vast majority of empirical research in economics envisions the possibility that demand curves can shift with changes in consumer tastes ($\varphi_{kt} \neq \varphi_{kt-1}$ for some $k \in \Omega$). Indeed, such shocks to expenditure shares conditional on prices (as captured by a structural residual in the demand system) are typically required in order for the model's predictions to be consistent with the observed data on prices and expenditure shares.

In contrast, our unified approach starts with the definition that the change in the cost of living between periods $t-1$ and t equals the cost of obtaining a unit of utility in period t divided by the cost of obtaining a unit of utility in period $t-1$:

$$\Phi_{t-1,t} = \frac{\mathbb{P}_t}{\mathbb{P}_{t-1}} = \frac{\mathbb{P}\left(\mathbf{P_t}, \boldsymbol{\varphi_t}, \sigma\right)}{\mathbb{P}\left(\mathbf{P_{t-1}}, \boldsymbol{\varphi_{t-1}}, \sigma\right)}. \tag{5}$$

We next use the demand system to substitute for the unobserved preference parameters ($\boldsymbol{\varphi_t}$, $\boldsymbol{\varphi_{t-1}}$) for each good and to estimate the substitution parameters (σ). In particular, using the invertibility of the demand system in equation (3), we express the unobserved demand parameters ($\boldsymbol{\varphi_t}$) in terms of observed prices and expenditure shares to obtain the following exact price index:

$$\Phi^{U}_{t-1,t} = \frac{\mathbb{P}_t}{\mathbb{P}_{t-1}} = \frac{\mathbb{P}\left(\mathbf{P_t}, \mathbb{S}^{-1}\left(\mathbf{S_t}, \mathbf{P_t}, \sigma, \bar{\varphi}\right), \sigma\right)}{\mathbb{P}\left(\mathbf{P_{t-1}}, \mathbb{S}^{-1}\left(\mathbf{S_{t-1}}, \mathbf{P_{t-1}}, \sigma, \bar{\varphi}\right), \sigma\right)}, \tag{6}$$

where the superscript U is a mnemonic for unified. We choose a common set of units in which to measure the demand parameters in each time period (any finite positive scalar $\bar{\varphi}_t = \bar{\varphi}$ for all t). Given these common units, the exact price index (6) is money-metric, because it depends solely on prices and expenditure shares that are directly comparable over time. Therefore, our unified approach simultaneously allows for demand shocks for individual goods in the demand system in equation (3), while preserving money-metric utility in the price index in equation (6), and hence permitting consistent comparisons on welfare over time.

Conventional price indices in equation (4) are only valid under the standard assumption that the demand parameters for *all* goods are constant over time ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega$). Our price index in equation (6) is also valid under this standard assumption, because we substitute out for the unobserved demand parameters in each period using the observed expenditure shares. The key difference between our approach and the standard approach is that our price index is also valid under a much weaker set of assumptions, namely that there are demand shocks for individual goods ($\varphi_{kt} \neq \varphi_{kt-1}$ for some $k \in \Omega$), but that the demand parameters are measured in common units across time periods (the scalar $\bar{\varphi}_t = \bar{\varphi}$ for all t). Hence our approach holds under a much wider set of assumptions about demand than standard index numbers. We also resolve a deep tension inherent in these existing measures. Conventional price indices use the observed data on prices and expenditure shares as inputs, but they cannot rationalize these observed inputs as equilibrium outcomes, because they cannot allow for the shifts in expenditure shares conditional on prices required to explain the observed data. Therefore, whenever there are shifts in demand curves in the observed data, but these are assumed away by standard price indices, this will introduce a bias into these conventional measures, because these indices provide no way of interpreting a change in expenditure shares without a change in relative prices. We provide a formal characterization of this bias for CES and translog preferences below.

Compared to the standard approach, we use more of the structure of the model, because we combine the unit expenditure function and demand system. We also make an additional assumption, namely that the demand system is invertible. But we show below that this assumption is satisfied for many functional forms considered in the price index literature (including CES and translog preferences). Additionally, our approach requires estimates of the substitution parameters (σ) from the demand system. However, we show below that estimates of these parameters are also required in the existing price index literature, as soon as one allows for the entry and exit of goods, as observed in micro data. Furthermore, we use our unified approach below to derive a new estimator of the elasticity of substitution for the case of CES preferences.

3 The Unified Price Index

To implement our approach and illustrate the quantitative magnitude of the bias in conventional price indices from abstracting from demand shocks, we now assume a particular functional form for preferences. We focus on one of the most commonly used preference structures in economics, constant elasticity of substitution (CES) preferences. We derive what we term the "unified price index," which is money-metric, allows for demand shocks for individual goods, incorporates the entry and exit of goods over time, and nests the main existing economic and statistical approaches to the measurement of changes in the cost of living.

3.1 Preferences and Demand

The unit expenditure function ($\mathbb{P}_t$) is defined over the price (P_{kt}) of each good k at time t:

$$\mathbb{P}_t = \left[\sum_{k \in \Omega_t} \left(\frac{P_{kt}}{\varphi_{kt}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}, \qquad \sigma > 1, \qquad \varphi_{kt} > 0. \tag{7}$$

where σ is the elasticity of substitution across goods; φ_{kt} is the preference ("demand") parameter for good k at time t; and the set of goods supplied at time t is denoted by Ω_t.[3] Although we allow demand parameters for individual goods (φ_{kt}) to change over time, we assume a constant elasticity of substitution (σ) over time, as is required for money-metric utility. Applying Shephard's Lemma to this unit expenditure function, we obtain the demand system in which the expenditure share (S_{kt}) for each good k and time period t is:

$$S_{kt} \equiv \frac{P_{kt} C_{kt}}{\sum_{\ell \in \Omega_t} P_{\ell t} C_{\ell t}} = \frac{\left(P_{kt} / \varphi_{kt} \right)^{1-\sigma}}{\sum_{\ell \in \Omega_t} \left(P_{\ell t} / \varphi_{\ell t} \right)^{1-\sigma}}, \qquad k \in \Omega_t. \tag{8}$$

where φ_{kt} is a structural residual that ensures that the model exactly replicates the observed data on prices and expenditure shares as an equilibrium. We interpret this structural residual as a shock to consumer preferences in our bar-code data, as discussed above.

To allow for the entry and exit of goods over time, we partition the set of goods in period t (Ω_t) into those "common" to t and $t-1$ ($\Omega_{t,t-1}$) and those added between $t-1$ and t (I_t^+), where $\Omega_t = \Omega_{t,t-1} \cup I_t^+$. Similarly, we partition the set of goods in period $t-1$ (Ω_{t-1}) into those common to t and $t-1$ ($\Omega_{t,t-1}$) and those dropped between $t-1$ and t (I_t^-), where $\Omega_{t-1} = \Omega_{t,t-1} \cup I_{t-1}^-$. We denote the number of goods in period t by $N_t = |\Omega_t|$ and the number of common goods by $N_{t,t-1} = |\Omega_{t,t-1}|$. We assume that $\varphi_{kt} = 0$ for a good k before it enters and after it exits, which rationalizes the observed entry and exit of goods over time.

[3] We focus on CES preferences as in Dixit and Stiglitz (1977) and abstract from the generalizations of the love of variety properties of CES in Benassy (1996) and Behrens et al. (2014).

3.2 Changes in the Cost of Living

We now combine the unit expenditure function (7) and demand system (8) to derive our unified price index, taking into account the entry and exit of goods and changes in demand for each good. We start by expressing the change in the cost of living from $t-1$ to t as the ratio between the unit expenditure functions (7) in the two periods:

$$\Phi_{t-1,t} = \frac{\mathbb{P}_t}{\mathbb{P}_{t-1}} = \left[\frac{\sum_{k\in\Omega_t}\left(P_{kt}/\varphi_{kt}\right)^{1-\sigma}}{\sum_{k\in\Omega_{t-1}}\left(P_{kt-1}/\varphi_{kt-1}\right)^{1-\sigma}}\right]^{\frac{1}{1-\sigma}}. \tag{9}$$

The fact that the set of goods is changing means that the set of goods in the denominator is not the same as that in the numerator. Feenstra (1994) showed that one way around this problem is to express the price index in terms of price index for "common goods" (i.e., goods available in both time periods) and a variety-adjustment term. Summing equation (8) over the set of commonly available goods, we can express expenditure on all common goods as a share of total expenditure in periods t and $t-1$ respectively as:

$$\lambda_{t,t-1} \equiv \frac{\sum_{k\in\Omega_{t,t-1}}\left(P_{kt}/\varphi_{kt}\right)^{1-\sigma}}{\sum_{k\in\Omega_t}\left(P_{kt}/\varphi_{kt}\right)^{1-\sigma}}, \qquad \lambda_{t-1,t} \equiv \frac{\sum_{k\in\Omega_{t,t-1}}\left(P_{kt-1}/\varphi_{kt-1}\right)^{1-\sigma}}{\sum_{k\in\Omega_{t-1}}\left(P_{kt-1}/\varphi_{kt-1}\right)^{1-\sigma}}, \tag{10}$$

where $\lambda_{t,t-1}$ is equal to the total sales of continuing goods in period t divided by the sales of all goods available in time t evaluated at *current* prices. Its maximum value is one if no goods enter in period t and will *fall* as the share of *new* goods *rises*. Similarly, $\lambda_{t-1,t}$ is equal to total sales of continuing goods as share of total sales of all goods in the *past* period evaluated at $t-1$ prices. It will equal one if no goods cease being sold and will *fall* as the share of *exiting* goods *rises*.

Multiplying the numerator and denominator of the fraction inside the square parentheses in (9) by the summation $\sum_{k\in\Omega_{t,t-1}}\left(P_{kt}/\varphi_{kt}\right)^{1-\sigma}$ over common goods at time t, and using the definition of $\lambda_{t,t-1}$ in (10), we obtain:

$$\Phi_{t-1,t} = \left[\frac{1}{\lambda_{t,t-1}}\frac{\sum_{k\in\Omega_{t,t-1}}\left(P_{kt}/\varphi_{kt}\right)^{1-\sigma}}{\sum_{k\in\Omega_{t-1}}\left(P_{kt-1}/\varphi_{kt-1}\right)^{1-\sigma}}\right]^{\frac{1}{1-\sigma}}.$$

Multiplying the numerator and denominator by the summation $\sum_{k\in\Omega_{t,t-1}}\left(P_{kt-1}/\varphi_{kt-1}\right)^{1-\sigma}$ over common goods at time $t-1$, and using the definition of $\lambda_{t-1,t}$ in (10), we obtain the exact CES price index:

$$\Phi_{t-1,t} = \left[\frac{\lambda_{t-1,t}}{\lambda_{t,t-1}}\frac{\sum_{k\in\Omega_{t,t-1}}\left(P_{kt}/\varphi_{kt}\right)^{1-\sigma}}{\sum_{k\in\Omega_{t,t-1}}\left(P_{kt-1}/\varphi_{kt-1}\right)^{1-\sigma}}\right]^{\frac{1}{1-\sigma}} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}}\frac{\mathbb{P}^*_t}{\mathbb{P}^*_{t-1}}, \tag{11}$$

where we use an asterisk to denote the value of a variable for the common set of goods (*i.e.*, goods available in periods t and $t-1$), such that $\mathbb{P}^*_t$ and $\mathbb{P}^*_{t-1}$ are the unit expenditure functions defined over the set of *common* goods in t and $t-1$ ($\Omega_{t,t-1}$):

$$\mathbb{P}^*_t \equiv \left[\sum_{k\in\Omega_{t,t-1}}\left(\frac{P_{kt}}{\varphi_{kt}}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{12}$$

The common goods price index ($\mathbb{P}^*_t/\mathbb{P}^*_{t-1}$) is the change in the cost of living if the set of goods is not changing, and it will prove to be a useful building block in our unified price index. The term multiplying it in equation (11) is the "variety-adjustment" term ($\left(\lambda_{t,t-1}/\lambda_{t-1,t}\right)^{1/(\sigma-1)}$). This term adjusts the common goods price index for entering and exiting goods. If new goods are more numerous than exiting goods or have lower demand-adjusted prices (i.e., lower (P_{kt}/φ_{kt})), then $\lambda_{t,t-1}/\lambda_{t-1,t} < 1$, and the price index ($\Phi_{t-1,t}$) will fall due to an increase in variety or the entering varieties being more appealing given their cost than the exiting varieties.

Having defined the shares of common goods in total expenditure in equation (10), we can also define the share of individual common good $k \in \Omega_{t,t-1}$ in expenditure on all common goods (S^*_{kt}):

$$S^*_{kt} \equiv \frac{P_{kt}C_{kt}}{\sum_{\ell \in \Omega_{t,t-1}} P_{\ell t}C_{\ell t}} = \frac{\left(P_{kt}/\varphi_{kt}\right)^{1-\sigma}}{\sum_{\ell \in \Omega_{t,t-1}} \left(P_{\ell t}/\varphi_{\ell t}\right)^{1-\sigma}}, \qquad k \in \Omega_{t,t-1}, \tag{13}$$

which takes the same form as the share of each good in total expenditure, except that the summation in the denominator is only over common goods.

We now combine equations (11), (12) and (13) to obtain an exact price index for the true change in the cost of living between periods $t-1$ and t:

$$\Phi_{t-1,t} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}} \prod_{k \in \Omega_{t,t-1}} \left(\frac{P_{kt}/\varphi_{kt}}{P_{kt-1}/\varphi_{kt-1}}\right)^{\omega^*_{kt}}, \qquad \omega^*_{kt} \equiv \frac{\frac{S^*_{kt}-S^*_{kt-1}}{\ln S^*_{kt}-\ln S^*_{kt-1}}}{\sum\limits_{\ell \in \Omega_{t,t-1}} \frac{S^*_{\ell t}-S^*_{\ell t-1}}{\ln S^*_{\ell t}-\ln S^*_{\ell t-1}}}, \tag{14}$$

as shown in Section A.1 of the web appendix. This exact price index (14) is a generalization of the Sato-Vartia price index (Sato 1976 and Vartia 1976), which corresponds to the special case in which demand is assumed to be time invariant for each good ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega_{t,t-1}$). The weights (ω^*_{kt}) correspond to a logarithmic mean of common goods expenditure shares in the two time periods and sum to one across common goods: $\sum_{k \in \Omega_{t,t-1}} \omega^*_{kt} = 1$.

To express the exact price index (14) in a money-metric form, we invert the demand system by dividing the expenditure share (13) by its geometric mean to obtain a solution for the demand parameter as a function of prices and expenditures:

$$\frac{\varphi_{kt}}{\tilde{\varphi}_t} = \frac{\varphi_{kt}}{\bar{\varphi}} = \frac{P_{kt}}{\tilde{P}_t}\left(\frac{S_{kt}}{\tilde{S}_t}\right)^{\frac{1}{\sigma-1}}, \tag{15}$$

where a tilde over a variable denotes a geometric average and the asterisk indicates that the geometric average is taken for the set of common goods, such that $\tilde{x}^*_t = \left(\prod_{k \in \Omega_{t,t-1}} x_{kt}\right)^{1/N_{t,t-1}}$ for the variable x_{kt}.

We choose common units in which to measure the demand parameters over time, which corresponds to the assumption that the demand parameters have a constant geometric mean ($\tilde{\varphi}_t = \bar{\varphi}$ for all t). Therefore, we allow demand shocks for individual goods, but we assume that these demand shocks are mean zero in logs: $\frac{1}{N_{t,t-1}} \sum_{k \in \Omega_{t,t-1}} \Delta \ln \varphi_{kt} = 0$. This condition necessarily holds under the standard assumption that the demand for each common good is time invariant ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega_{t,t-1}$). But it also holds under the much weaker assumption that demand shocks average out across common goods: $\frac{1}{N_{t,t-1}} \sum_{k \in \Omega_{t,t-1}} \Delta \ln \varphi_{kt} = 0$. This assumption is a necessary and sufficient condition for there to exist a common set of units in which to measure utility over time, and hence for a money-metric formulation of the price index to exist. One way of interpreting this assumption is that there is a constant underlying preference structure, but that each period there are random shocks to consumer tastes for each good that are drawn from a distribution with a constant mean.

We now derive our money-metric unified price index by using the inversion of the demand system in equation (15) to substitute for the demand parameters (φ_{kt}) in the exact price index in equation (14).[4]

Proposition 1. *The "unified price index" (UPI)—which is exact for the CES preference structure in the presence of changes in the set of goods, demand-shocks for individual goods, and discrete changes in prices and expenditure shares—is given by*

$$\Phi^{U}_{t-1,t} = \underbrace{\left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}}}_{\text{Variety Adjustment}} \underbrace{\left[\frac{\tilde{P}^*_t}{\tilde{P}^*_{t-1}}\left(\frac{\tilde{S}^*_t}{\tilde{S}^*_{t-1}}\right)^{\frac{1}{\sigma-1}}\right]}_{\text{Common-Goods Price Index } \Phi^{CG}_{t-t,t}} . \tag{16}$$

Proof. The proposition follows directly from substituting equation (15) into equation (14). □

The UPI has important similarities with other price indexes that enable us to nest many existing approaches. For example, as in Feenstra (1994), the unified price index (UPI) expresses the change in the cost of living as a function of a variety-adjustment term and a common-goods component of the unified price index (CG-UPI). The variety-adjustment term (namely $\left(\lambda_{t,t-1}/\lambda_{t-1,t}\right)^{1/(\sigma-1)}$ in equation (16)) captures changes in the unit expenditure function due to product turnover, changes in the number of varieties, and new goods. The CG-UPI (denoted by $\Phi^{CG}_{t-t,t}$ in equation (16)) measures how changes in prices, demand shifts, and product substitution for common goods affect a consumer's unit expenditure function. It is comprised of two terms. The first term ($\tilde{P}^*_t/\tilde{P}^*_{t-1}$) is none other than the geometric average of price relatives that serves as the basis for lower level of the U.S. Consumer Price Index (also known as the "Jevons" index). Indeed, in the special case in which varieties are perfect substitutes ($\sigma \to \infty$), the UPI collapses to the Jevons index, since both $\left(\lambda_{t,t-1}/\lambda_{t-1,t}\right)^{1/(\sigma-1)}$ and $\left(\tilde{S}^*_t/\tilde{S}^*_{t-1}\right)^{1/(\sigma-1)}$ converge to one as $\sigma \to \infty$.

The last term ($\left(\tilde{S}^*_t/\tilde{S}^*_{t-1}\right)^{1/(\sigma-1)}$) is novel and captures heterogeneity in expenditure shares across common goods. This term moves with the ratio of the geometric mean of common goods expenditure shares in the two periods. Critically, as the market shares of common goods in a time period become more uneven, the geometric average will fall. Thus, this term implies that the cost of living will fall if expenditure shares become more dispersed. Intuitively, when varieties are substitutes ($\sigma > 1$), consumers value dispersion in demand-adjusted prices across varieties, because they can substitute consumption towards more appealing varieties.[5]

The UPI in (16) has a number of desirable economic and statistical properties. First, this price index and each of its components are "time reversible" for any value of σ, thereby permitting consistent comparisons of welfare going forwards and backwards in time. In other words, given any set of product turnover, price

[4] Equivalently, our unified price index can be derived using the common goods expenditure share. Rewriting equation (13), we obtain: $P^*_t = \left(P_{kt}/\varphi_{kt}\right)\left(S^*_{kt}\right)^{1/(\sigma-1)}$. Taking geometric means of both sides of this equation, differencing over time, and using our common choice of units to measure the demand parameters ($\tilde{\varphi}_t = \bar{\varphi}$), we obtain the common goods price index component of equation (16): $\Phi^{CG}_{t-t,t} = \left(\tilde{P}^*_t/\tilde{P}^*_{t-1}\right)\left(\tilde{S}^*_t/\tilde{S}^*_{t-1}\right)^{1/(\sigma-1)}$.

[5] Our unified price index (16) differs from the expression for the CES price index in Hottman et al. (2016), which did not distinguish entering and exiting goods from common goods (omitting $\left(\lambda_{t,t-1}/\lambda_{t-1,t}\right)^{1/(\sigma-1)}$) and captured the dispersion of sales across common goods in different way (using a different term from $\left(\tilde{S}^*_t/\tilde{S}^*_{t-1}\right)^{1/(\sigma-1)}$).

changes, and demand shifts between $t-1$ and t, the percent change in prices between $t-1$ and t is the inverse of the change between t and $t-1$ (i.e., $\Phi^U_{t-1,t} = 1/\Phi^U_{t,t-1}$). Second, the unified price index depends in a simple and transparent way on the elasticity of substitution. Variation in this elasticity leaves the terms in common goods prices unchanged ($\tilde{P}^*_t / \tilde{P}^*_{t-1}$) and affects the variety adjustment $\left(\lambda_{t,t-1}/\lambda_{t-1,t}\right)^{1/(\sigma-1)}$) and heterogeneity terms ($\left(\tilde{S}^*_t / \tilde{S}^*_{t-1}\right)^{1/(\sigma-1)}$) depending on the extent to which these two expenditure share ratios are greater than or less than one. Finally, the relative magnitude of these variety and heterogeneity corrections in logs is independent of the value of the elasticity of substitution, and depends solely on the relative values of expenditure share moments in the data ($\ln\left((\lambda_{t,t-1}/\lambda_{t-1,t}\right) / \ln\left(\tilde{S}^*_t/\tilde{S}^*_{t-1}\right))$).

4 Relation to Existing Price Indexes

In this section, we compare our unified price index with all of the main economic and statistical price indexes used in the existing theoretical and empirical literature on price measurement. We first discuss the relationship between our index and other indexes for the CES demand system. We next show that all other conventional price indexes are special cases of the unified price index that either impose particular parameter restrictions (on the elasticity of substitution), abstract from the entry and exit of goods, and/or neglect changes in demand for each good.

4.1 Relation to Existing Exact CES Price Indexes

The formula for the UPI differs from the CES price index in Feenstra (1994) because we do not use the Sato (1976) and Vartia (1976) formula for the common goods price index. The Feenstra index is given by:[6]

$$\frac{\mathbb{P}^*_t}{\mathbb{P}^*_{t-1}} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}} \Phi^{SV}_{t-1,t}, \qquad \Phi^{SV}_{t-1,t} \equiv \prod_{k\in\Omega_{t,t-1}} \left(\frac{P_{kt}}{P_{kt-1}}\right)^{\omega^*_{kt}}, \qquad \omega^*_{kt} \equiv \frac{\frac{S^*_{kt}-S^*_{kt-1}}{\ln S^*_{kt}-\ln S^*_{kt-1}}}{\sum_{\ell\in\Omega_{t,t-1}} \frac{S^*_{\ell t}-S^*_{\ell t-1}}{\ln S^*_{\ell t}-\ln S^*_{\ell t-1}}}. \tag{17}$$

Both indexes require the estimation of σ, but our approach resolves a tension that Feenstra (1994) observed was inherent in his use of the Sato-Vartia formula. The Sato-Vartia index ($\Phi^{SV}_{t-1,t}$) used for $\mathbb{P}^*_t/\mathbb{P}^*_{t-1}$ assumes that demand is constant over time for each good ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega_{t,t-1}$ and t), whereas the estimation of σ assumes that demand for goods changes over time ($\varphi_{kt} \neq \varphi_{kt-1}$ for some k and t).

This tension is more pernicious than it might appear because the assumption of time-invariant demand is a *crucial* assumption for the derivation of the Sato-Vartia index, and the index cannot be derived if one assumes mean-zero log demand shocks. Under the assumption of constant demand for each common good ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega_{t,t-1}$), we show in the proposition below that there is no need to estimate σ, because it can be recovered from observed prices and expenditure shares using the weights from the Sato-Vartia price index. Furthermore, the model is overidentified when demand is constant for each common good, with the result that there exists an *infinite* number of approaches to measuring σ. If demand is indeed constant for each common good ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega_{t,t-1}$), each of these approaches returns exactly the

[6] As shown in Banerjee (1983), the Sato-Vartia weights (ω^*_{kt}) are only one of a broader class of weights that can be used to construct the exact common-goods CES price index with constant demand for each common good ($\varphi_{kt} = \varphi_k$).

same value for σ. However, if demand for goods changes over time ($\varphi_{kt} \neq \varphi_{kt-1}$ for some $k \in \Omega_{t,t-1}$), and a researcher falsely assumes constant demand for each good, we show that each of these approaches returns a different value for σ in every time period. Even making the additional assumption that on average the change in demand for goods is zero for common goods does not eliminate the problem. These approaches produce a different value for σ unless demand is constant for every common good.

Proposition 2. ***(a)*** *Under the assumption that demand is constant for each common good ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega_{t,t-1}$ and t), the elasticity of substitution (σ) is uniquely identified from observed changes in prices and expenditure shares with no estimation. Furthermore, there exists a continuum of approaches to measuring σ, each of which weights prices and expenditure shares with different non-negative weights that sum to one, but returns the same value for σ.*
(b) *If demand for common goods changes over time ($\varphi_{kt} \neq \varphi_{kt-1}$ for some $k \in \Omega_{t,t-1}$ and t), but a researcher falsely assumes that demand for each common good is constant, each of these alternative approaches returns a different value for σ, depending on which non-negative weights are used.*

Proof. See Section A.2 of the web appendix. □

This proposition makes clear the link between the common-goods component of the unified price index and the standard Sato-Vartia CES price index. If there are no demand shifts, the two indexes are identical. In the presence of non-zero demand shifts, the CG-UPI exactly replicates the observed data on expenditure shares and prices as an equilibrium of the model based on the assumption of a constant elasticity of substitution (σ) and time-varying demand (φ_{kt}). In contrast, the Sato-Vartia index assumes time-invariant demand for each good, which implies that the model does not exactly replicate the observed data on expenditure shares and prices if there are non-zero demand shifts. The elasticity of substitution implied by the Sato-Vartia index will vary with these demand shifts, which makes the Sato-Vartia price index depend on demand parameters and therefore incompatible with a money-metric utility function. The implicit elasticity of substitution in the Sato-Vartia CES price index is is not only time varying (a property we will explore in Section 7.2), but also will differ based on what arbitrary subset of common goods are included in the index and how one weights them. Therefore, if there are demand shifts, standard CES price indexes imply that the elasticity of substitution is not constant within a time period or across them, rendering the utility function time varying and traditional welfare analysis problematic. By contrast, a key advantage of the UPI is that it results in a money-metric utility function even in the presence of these demand shocks for individual goods.

This problem also biases any attempt to measure aggregate price changes using a Sato-Vartia formula in the presence of demand shifts as the following proposition demonstrates.

Proposition 3. *In the presence of non-zero demand shocks for some good (i.e.,* $\ln\left(\varphi_{kt}/\varphi_{kt-1}\right) \neq 0$ *for some $k \in \Omega_{t,t-1}$), the Sato-Vartia price index ($\Phi^{SV}_{t-1,t}$) differs from the exact common goods CES price index. The Sato-Vartia price index ($\Phi^{SV}_{t-1,t}$) equals the unified price index (16) plus a demand shock bias term.*

$$\ln \Phi^{SV}_{t-1,t} = \ln \Phi^{CG}_{t-1,t} + \underbrace{\left[\sum_{k \in \Omega_{t,t-1}} \omega^*_{kt} \ln \left(\frac{\varphi_{kt}}{\varphi_{kt-1}} \right) \right]}_{\textit{demand shock bias}}, \tag{18}$$

$$\textit{where} \qquad \varphi_{kt} = \frac{P_{kt}}{\tilde{P}^*_t} \left(\frac{S_{kt}}{\tilde{S}^*_t} \right)^{\frac{1}{\sigma-1}}, \qquad \omega^*_{kt} \equiv \frac{\frac{S^*_{kt} - S^*_{kt-1}}{\ln S^*_{kt} - \ln S^*_{kt-1}}}{\sum_{\ell \in \Omega_{t,t-1}} \frac{S^*_{\ell t} - S^*_{\ell t-1}}{\ln S^*_{\ell t} - \ln S^*_{\ell t-1}}}, \qquad \sum_{k \in \Omega_{t,t-1}} \omega^*_{kt} = 1. \tag{19}$$

Proof. See Section A.3 of the web appendix. □

While the Sato-Vartia index computes the expenditure-share weighted average of price changes, the true cost of living depends on the expenditure-share weighted average of *demand-adjusted* price changes. Therefore, the Sato-Vartia index will be unbiased if the demand shocks ($\ln(\varphi_{kt} / \varphi_{kt-1})$) are orthogonal to the expenditure-share weights (ω^*_{kt}), upward-biased if they are positively correlated with these weights, and downward-biased if they are negatively correlated with these weights. This bias arises even though the demand shocks are mean zero in logs. The intuition is as follows. If the demand shocks are positively correlated with the expenditure-share weights, the demand for goods that account for large shares of the consumer's cost of living rises relatively more than for goods that account for small shares, which reduces the weighted-average of demand-adjusted prices, and hence reduces the cost of living. Conversely, if the demand shocks are negatively correlated with the expenditure-share weights, the opposite is true, which increases the weighted-average of demand-adjusted prices, and hence raises the cost of living.

In principle, either a positive or negative correlation between the demand shocks ($\ln(\varphi_{kt} / \varphi_{kt-1})$) and the expenditure-share weights (ω^*_{kt}) is possible, depending on the underlying correlation between demands (φ_{kt}) and prices (P_{kt}) in the two time periods. However, there is a mechanical force for a positive correlation, because the expenditure-share weights themselves are functions of the demand shocks. In particular, as shown in the proposition below, a positive demand shock for a good mechanically increases the expenditure-share weight for that good and reduces the expenditure-share weight for all other goods.

Proposition 4. *A positive demand shock for a good k (i.e., $\ln(\varphi_{kt} / \varphi_{kt-1}) > 0$ for some $k \in \Omega_{t,t-1}$) increases the Sato-Vartia weight for that good (ω^*_{kt}) and reduces the Sato-Vartia weight for all other goods $\ell \neq k$ ($\omega^*_{\ell t}$).*

Proof. See Section A.4 of the web appendix. □

Intuitively, if a consumer starts to value a good more in period t relative to $t-1$ ($\varphi_{kt} > \varphi_{kt-1}$), she substitutes towards that good ($S^*_{kt} > S^*_{kt-1}$), which raises its expenditure-share weight (ω^*_{kt}), other things equal. We therefore refer to this bias as the "consumer-valuation" bias, because it reflects the fact that consumers can substitute towards the goods that they value more. Since demand (φ_{kt}) enters inversely with prices (P_{kt}), this consumer-valuation bias is closely related to the "substitution bias" in fixed-weight price indices, which do not take into account that consumers can substitute towards the goods for which prices fall. We provide empirical evidence on this consumer-valuation bias below.

As discussed above, we focus on CES preferences, because of their prevalence in economics, and because they provide a tractable approach to incorporating entry and exit and estimating the substitution parameter between goods. In Section A.16 of the web appendix, we implement our unified approach for homothetic translog preferences with a constant set of goods (Ω). The translog functional form is flexible in the sense that it provides a second-order approximation to any twice continuously differentiable expenditure function. Analogous to Proposition 3, we show that the Törnqvist index is upward biased if demand shocks are positively correlated with the arithmetic mean of expenditure shares in the two time periods. Similar to Proposition 4, we show that there is a mechanical force that tends to generate such a positive correlation, because the expenditure share at time t (S_{kt}^{*}) is endogenous to the demand shock ($\varphi_{kt}/\varphi_{kt-1}$). Therefore, we find that conventional price indices are also subject to a consumer-valuation bias.

Our use of bar-code data implies that a change in the structural residual (φ_{kt}) cannot correspond to a change in product quality, because firms have strong incentives of inventory and stock control not to use the same bar code for products with different observable characteristics. Similarly, our use of bar-code data alleviates concerns about measurement error. Nonetheless, a further advantage of our exact common-goods price index in other contexts is that it is invariant with respect to mean zero log additive measurement error in prices and expenditure shares. In contrast, the measured Sato-Vartia common goods price index in equation (17) in general differs from its true value in the presence of such measurement error.

Proposition 5. *The common-goods unified price index price index ($\Phi_{t-1,t}^{CG-UPI}$) in equation (16) is invariant with respect to mean-zero log additive measurement error in either prices and/or expenditure shares. In contrast, the measured Sato-Vartia common goods price index ($\Phi_{t-1,t}^{SV}$) in equation (17) in general differs from its true value in the presence of such measurement error.*

Proof. See Section A.5 of the web appendix. □

In conclusion, Propositions 2-4 show that there are three major differences between our index (16) and the Feenstra index. First, if one assumes that demand for each good is time invariant when it is in fact time varying, the Sato-Vartia formula arbitrarily implies one of an infinite set of elasticities that are consistent with the CES functional form, and none of these need be consistent with the elasticity identified using econometric techniques. Thus, our index eliminates the inconsistency that Feenstra (1994) identified as arising from imposing no demand shocks when computing the price change for the common goods component of the CES price index while also assuming these shocks to be time varying when estimating σ for the variety correction term ($(\lambda_{t,t-1}/\lambda_{t-1,t})^{1/(\sigma-1)}$). Second, we show that the assumption of time-invariant demand in the construction of price indexes introduces an upward "consumer-valuation bias," because of the counterfactual assumption that consumers will not shift expenditures towards goods they prefer. Third, our expression for the exact common goods component of the CES price index has the desirable property that it is robust to mean zero log additive measurement error in either prices and/or expenditure shares.

4.2 Relation to Conventional Price Indexes

Our unified price index is exact for the CES functional form and expresses changes in the cost of living solely in terms of prices and expenditure shares. However, there are two other equivalent expressions for the change in the cost of living in terms of prices, expenditure shares and demands for each good that are not necessarily money metric. These equivalent expressions arise from forward and backward differences of the unit expenditure function and we now make them explicit in order to relate our approach to other conventional price indexes and to later show how our approach can be used to estimate the elasticity of substitution between goods.

The forward difference of the unit expenditure function evaluates the *increase* in the price index from $t-1$ to t using the expenditure shares of consumers in period $t-1$. Using equations (10), (11), (12) and (13), this forward difference can be written in terms of the change in variety ($\lambda_{t,t-1}/\lambda_{t-1,t}$), the initial share of each common good in expenditure on all common goods (S^*_{kt-1}), and changes in prices (P_{kt}/P_{kt-1}) and demand ($\varphi_{kt}/\varphi_{kt-1}$) for all common goods:

$$\Phi^F_{t-1,t} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}} \frac{\mathbb{P}^*_t}{\mathbb{P}^*_{t-1}} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}} \left[\sum_{k\in\Omega_{t,t-1}} S^*_{kt-1}\left(\frac{P_{kt}/\varphi_{kt}}{P_{kt-1}/\varphi_{kt-1}}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}, \tag{20}$$

as shown in Section A.6 of the web appendix. The backward difference of the unit expenditure function uses the expenditure shares of consumers period t to evaluate the *decrease* in the price index from t to $t-1$. Using equations (10), (11), (12) and (13), this backward difference can be written in an analogous form as:

$$\Phi^B_{t,t-1} = \left(\frac{\lambda_{t-1,t}}{\lambda_{t,t-1}}\right)^{\frac{1}{\sigma-1}} \frac{\mathbb{P}^*_{t-1}}{\mathbb{P}^*_t} = \left(\frac{\lambda_{t-1,t}}{\lambda_{t,t-1}}\right)^{\frac{1}{\sigma-1}} \left[\sum_{k\in\Omega_{t,t-1}} S^*_{kt}\left(\frac{P_{kt-1}/\varphi_{kt-1}}{P_{kt}/\varphi_{kt}}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}, \tag{21}$$

where the algebra is again relegated to Section A.6 of the web appendix.[7]

The only variable not in common to the forward and backward differences is the expenditure share (S^*_{kt-1} versus S^*_{kt}). When evaluating the change in the cost of living going forward in time, we use the period $t-1$ expenditure shares, whereas when doing the same going backward in time, we use the period t expenditure shares. The terms in square brackets in (16), (20) and (21) correspond to three equivalent ways of expressing the change in the cost of living for common goods. We now use these three equivalent expressions to show that all conventional price indexes correspond to special cases of our unified price index that impose particular parameter restrictions, abstract from changes in demand for each good, and/or abstract from the entry and exit of goods over time.

According to an International Labor Organization (ILO) survey of 68 countries around the world, the Dutot (1738) index is still the most prominent one for measuring price changes (Stoevska (2008)).[8] This index is the ratio of a simple average of prices in two periods:

[7] The forward and backward differences in equations (20) and (21) are related to the comparisons of welfare using initial and final preferences considered in Fisher and Shell (1972). A key difference is that our expressions (20) and (21) include the change in demand for each good ($\varphi_{kt}/\varphi_{kt-1}$), and hence are exactly equal to the unified price index (16), rather than providing bounds for it.

[8] 41 percent of countries use this index although historically its popularity was much higher. For example, all U.S. inflation data prior to 1999 is based on this index, and Belgian, German, and Japanese data continues to be based on it. The ILO report can be accessed here: http://www.ilo.org/public/english/bureau/stat/download/cpi/survey.pdf

$$\Phi^D_{t-1,t} \equiv \frac{\frac{1}{N_{t,t-1}}\sum_{k\in\Omega_{t,t-1}} P_{kt}}{\frac{1}{N_{t,t-1}}\sum_{k\in\Omega_{t,t-1}} P_{kt-1}} = \sum_{k\in\Omega_{t,t-1}} \frac{P_{k,t-1}}{\sum_{k\in\Omega_{t,t-1}} P_{kt-1}} \left(\frac{P_{kt}}{P_{k,t-1}}\right) \tag{22}$$

As the above formula shows, this index is simply a price-weighted average change in prices, which does not have a clear rationale in terms of economic theory.

A price-weighted average of price changes is a sufficiently problematic way of measuring changes in the cost of living that most statistical agencies do not just compute unweighted averages of prices in two periods, but select their sample of price quotes based on the largest selling products in the first period. If we think that the probability that a statistical agency picks a product for inclusion in its sample of prices is based on its purchase frequency ($C_{\ell,t-1}/\sum_{k\in\Omega_{t-1,t}} C_{k,t-1}$), then the Dutot index, as it is typically implemented, becomes the more familiar Laspeyres index, as used in U.S. import and export price indices:

$$\Phi^L_{t-1,t} \equiv \frac{\sum_{k\in\Omega_{t,t-1}} C_{k,t-1}P_{kt}}{\sum_{k\in\Omega_{t,t-1}} C_{k,t-1}P_{kt-1}} = \sum_{k\in\Omega_{t,t-1}} \frac{C_{k,t-1}P_{k,t-1}}{\sum_{k\in\Omega_{t,t-1}} C_{k,t-1}P_{kt-1}} \left(\frac{P_{kt}}{P_{k,t-1}}\right) = \sum_{k\in\Omega_{t,t-1}} S^*_{kt-1}\frac{P_{kt}}{P_{kt-1}}. \tag{23}$$

Written this way, it is clear that the Laspeyres index is a special case of our CES price index (20) in which the utility gain of new goods is exactly offset by the loss from disappearing goods ($\lambda_{t,t-1}/\lambda_{t-1,t} = 1$), the elasticity of substitution equals zero and demand for each good is constant ($\varphi_{kt}/\varphi_{kt-1} = 1$).

The Carli index, used by 19 percent of countries, is another popular index that can be thought of as a variant of the Laspeyres index. The formula for the Carli index is

$$\Phi^C_{t-1,t} \equiv \sum_{k\in\Omega_{t,t-1}} \frac{1}{N_{t,t-1}} \left(\frac{P_{kt}}{P_{k,t-1}}\right) \tag{24}$$

This index is identical to the Laspeyres if all goods have equal expenditure shares. However, as with the Dutot, it is important to remember that statistical agencies are more likely to select a good for inclusion in the sample if it has a high sales share ($S^*_{k,t-1}$). In this case, the Carli index also collapses back to the Laspeyres formula.

Similarly, the Paasche index is closely related to the Laspeyres index with the only difference that it weights price changes from $t-1$ to t by their expenditure shares in the *end* period t:

$$\Phi^P_{t-1,t} = \frac{\sum_{k\in\Omega_{t,t-1}} P_{kt}C_{kt}}{\sum_{k\in\Omega_{t,t-1}} P_{kt-1}C_{kt}} = \left[\sum_{k\in\Omega_{t,t-1}} S^*_{kt}\left(\frac{P_{kt}}{P_{kt-1}}\right)^{-1}\right]^{-1}. \tag{25}$$

We can also think of the Paasche index as is a special case of the CES price index (21) in which we apply the same parameter restrictions to derive the Laspeyres index.[9]

Finally, the Jevons index, which forms the basis of the lower level of the U.S. Consumer Price Index, is the second-most popular index currently in use, with 37 percent of countries building their measures of changes in the cost of living based on it.[10] The index is constructed by taking an unweighted geometric mean of price

[9] To derive (25) from (21), we use $\Phi_{t-1,t} = 1/\Phi_{t,t-1}$, assume $\lambda_{t-1,t}/\lambda_{t,t-1} = 1$ and $\varphi_{kt}/\varphi_{kt-1} = 1$ for all k, and set $\sigma = 0$.

[10] The percentages do not sum to 100 because 3 percent of sample respondents used other formulas.

changes from $t-1$ to t:

$$\Phi_{t-1,t}^{J} = \prod_{k\in\Omega_{t,t-1}} \left(\frac{P_{kt}}{P_{kt-1}}\right)^{\frac{1}{N_{t,t-1}}} = \frac{\tilde{P}_t^*}{\tilde{P}_{t-1}^*}. \tag{26}$$

As we discussed earlier, this formula is just a special case of the unified price index (16) in the limit as $\sigma \to \infty$. It is also related to the unified price index through another route. Statistical agencies typically choose products based on their historic sales shares. In this case the Jevons index becomes:

$$\Phi_{t-1,t}^{CD} = \prod_{k\in\Omega_{t,t-1}} \left(\frac{P_{kt}}{P_{kt-1}}\right)^{S_{k,t-1}^*}, \tag{27}$$

which Konyus (Konüs) and Byushgens (1926) proved was exact for the Cobb-Douglas (1928) functional form. This price index is a special case of the CES price index when the elasticity of substitution equals one, demand for each good is constant, and there are no changes in variety.

Existing measures of changes in the cost of living are therefore special cases of the unified approach developed in this paper, and biases can be thought of in terms of parameter restrictions on the unified price index. For example, "substitution bias" arises from building a price index using the wrong elasticity of substitution (σ). Most studies of consumer behavior suggest that this elasticity is greater than one, but in Laspeyres and Paasche indexes it arises because this elasticity is assumed to be zero. The recent move to the Jevons index by many countries reduced the substitution bias by changing the elasticity in the unified price index to infinity or, if one reinterprets the Jevons index as a Cobb-Douglas index, an elasticity of one. Our index corrects for this shortcoming in previous indexes by letting the data determine the correct elasticity.

"Variety" or "New Goods Bias" arises from the assumption that $\lambda_{t,t-1}/\lambda_{t-1,t} = 1$, which means that the utility gain from new goods is exactly offset by the loss from disappearing goods.[11] The fact that quite often the price per unit quality of new goods is lower than that of disappearing goods—for example, a \$1,000 computer is better today than ten years ago—implies that conventional price indexes are biased upwards because $\lambda_{t,t-1}/\lambda_{t-1,t} < 1$. In contrast, our index explicitly incorporates new and disappearing goods into the measurement of changes in the cost of living.

The third "consumer-valuation" bias is novel and arises because of the assumption that consumer demand for each good is constant over time ($\varphi_{kt}/\varphi_{kt-1} = 1$). Mechanically, this arises whenever a price index specifies that prices should be deflated by a demand parameter that is time varying (i.e., when the unit expenditure function depends on P_{kt}/φ_{kt}). In this sense, it is isomorphic to the well-known substitution bias that plagues fixed-weight indexes like the Laspeyres. Analogously, the consumer-valuation bias arises whenever one fixes the utility parameter associated with a good because it assumes consumers will not change expenditure patterns when their tastes change.

Interestingly, the two remaining "superlative" price indexes (Fisher and Törnqvist) are also closely related to the CES. Taking the geometric mean of the forward and backward differences of the CES price index (20) and (21), which are equal to the unified price index (16), we obtain the following quadratic mean of order

[11] The new goods bias is typically stated in terms of an index not allowing for new goods, but this is not technically correct. The absence of new goods would correspond to $\lambda_{t,t-1} = 1$. While it is true that if there are no new or exiting goods, we will have $\lambda_{t,t-1} = \lambda_{t-1,t} = 1$, the validity of common goods price indexes depends on a slightly weaker assumption: $\lambda_{t,t-1}/\lambda_{t-1,t} = 1$.

$2(1-\sigma)$ price index (Diewert 1976):

$$\Phi_{t,t-1}=\left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}}\left[\frac{\sum_{k\in\Omega_{t,t-1}}S_{kt-1}^{*}\left(\frac{P_{kt}/\varphi_{kt}}{P_{kt-1}/\varphi_{kt-1}}\right)^{1-\sigma}}{\sum_{k\in\Omega_{t,t-1}}S_{kt}^{*}\left(\frac{P_{kt}/\varphi_{kt}}{P_{kt-1}/\varphi_{kt-1}}\right)^{-(1-\sigma)}}\right]^{\frac{1}{2(1-\sigma)}}, \tag{28}$$

The Fisher index is the geometric mean of the Laspeyres (23) and Paasche (25) price indexes, and corresponds to the special case of (28) in which $\sigma=0$, the utility gain from new goods is exactly offset by the loss from disappearing goods ($\lambda_{t,t-1}/\lambda_{t-1,t}=1$), and demand for each good is constant ($\varphi_{kt}/\varphi_{kt-1}=1$):

$$\Phi_{t-1,t}^{F}=\left(\Phi_{t-1,t}^{L}\Phi_{t-1,t}^{P}\right)^{1/2}. \tag{29}$$

Closely related to the Fisher index is the Törnqvist index, which corresponds to the limiting case of (28) in which $\sigma\to 1$, the utility gain from new goods is exactly offset by the loss from disappearing goods ($\lambda_{t,t-1}/\lambda_{t-1,t}=1$), and demand for each good is constant ($\varphi_{kt}/\varphi_{kt-1}=1$):

$$\Phi_{t-1,t}^{T}=\prod_{k\in\Omega_{t,t-1}}\left(\frac{P_{kt}}{P_{kt-1}}\right)^{\frac{1}{2}\left(S_{kt-1}^{*}+S_{kt}^{*}\right)}. \tag{30}$$

Another way of looking at the Törnqvist index is to realize that it is just a geometric average of Cobb-Douglas price indexes defined in equation (27) evaluated at times $t-1$ and t.

The Fisher and Törnqvist price indexes are exact in the sense that they hold for flexible functional forms: quadratic mean of order-r preferences and the translog expenditure function respectively (Diewert 1976). These price indexes are also superlative in the sense that they provide a local second-order approximation to any continuous and differentiable expenditure function. However, we have shown that both indexes are closely related to the CES price index, and are in fact special cases of the geometric mean of two of our equivalent expressions for the CES price index (28) for a particular value of the elasticity of substitution. Therefore, the CES, Fisher, and Törnqvist price indexes for common goods are all closely related functions of the same underlying price and expenditure data. Empirically, we show below that the differences between these three indexes are trivially small when measured using only common goods and under the assumption of no demand shifts. Importantly, the exact and superlative properties of the Fisher and Törnqvist indexes are derived under the assumption of no entry and exit of goods and no changes in demand for each good. A key advantage of our unified price index (16) relative to these other two price indexes is that it explicitly takes into account both product turnover and changes in consumer valuations of each good, which we show below to be central features of micro data on prices and expenditure shares.[12]

[12] As discussed above, in Section A.16 of the web appendix, we implement our unified approach for translog preferences, and show that analogs of Propositions 3 and 4 hold for these preferences. In particular, the exact price index for translog preferences differs from the Törnqvist index in the presence of demand shocks for individual goods. In Section A.17 of the web appendix, we show that continuous Divisia indices are also subject to the consumer-valuation bias, because they abstract from changes in demand for individual goods.

Figure 1: Relation Between Existing Indexes and the UPI

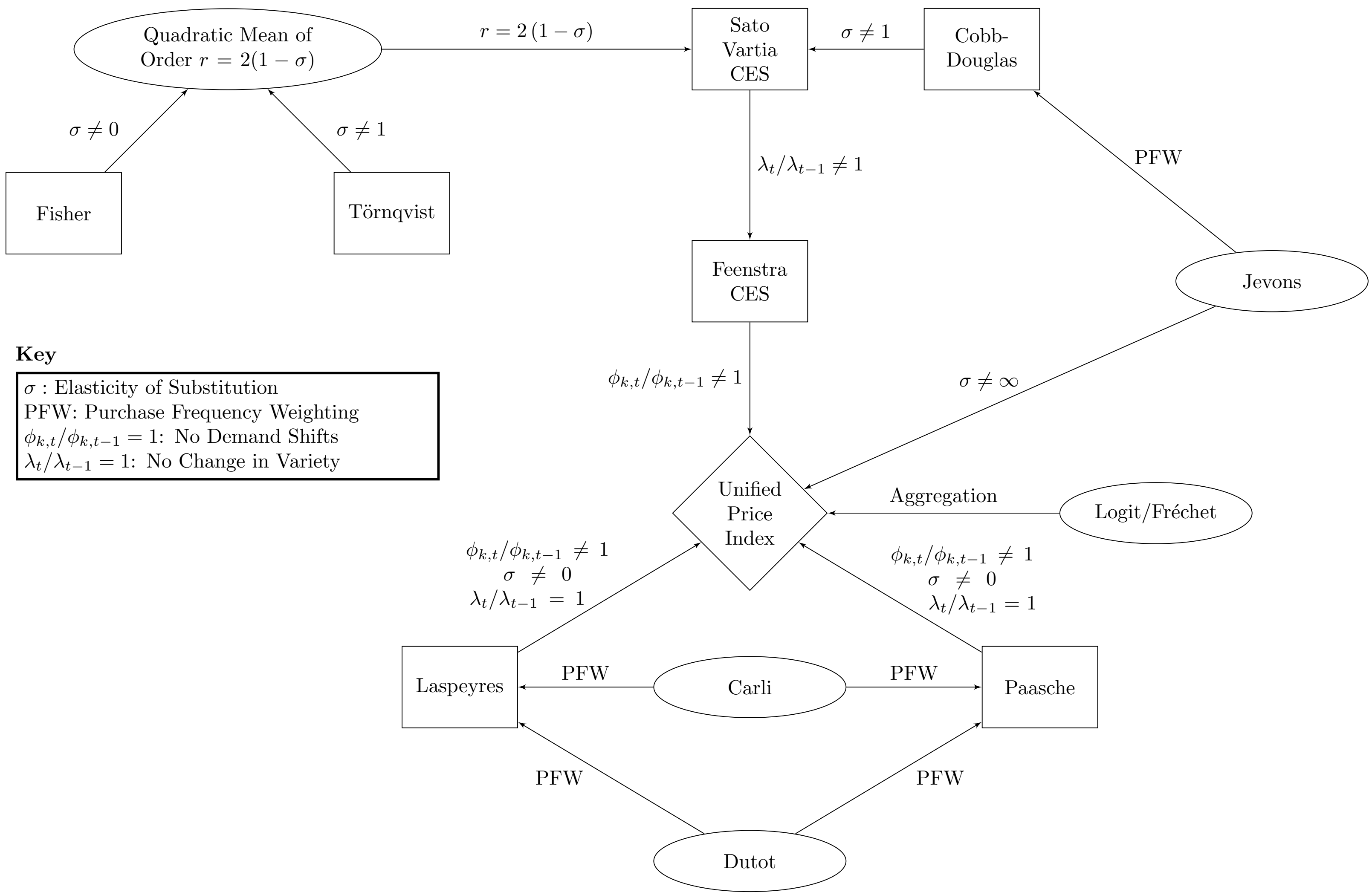


Figure 1 summarizes how all major price indexes are related to our unified index. Most existing indexes (such as the Dutot, Carli, Laspeyres, Paasche, Jevons, Cobb-Douglas, Sato-Vartia-CES, Feenstra-CES) are simply special cases of our index. Therefore, one can think of the standard approach to index numbers as versions of the unified approach, in which researchers make different parameter restrictions, ignore certain parts of the data (e.g., new goods), ignore certain implications of the model (e.g., the demand system and the unit expenditure function are derived from the same preferences), and fail to sample based on purchase frequencies. Existing exact CES price indexes assume no demand shocks, and superlative indexes are simply different weighted averages of the same building blocks as those of the unified index under the assumption of no change in the set of goods or the demand parameters for individual goods. The relaxation of all of these assumptions and restrictions results in the unified approach.

5 The UPI with Heterogeneous Consumers

In this section, we show that our unified price index encompasses versions of random utility models, such as Fréchet and logit, as examined in the discrete choice literature, including McFadden (1974), Anderson, de Palma, and Thisse (1992), and Nevo (2003). In these random utility models, the aggregate behavior of consumers with idiosyncratic tastes is the same as that of a representative consumer with CES preferences. Therefore, our unified price index captures the expected change in the cost of living for each consumer prior to the realization of idiosyncratic tastes. We also use this section to develop an extension to multiple types of consumers that have different substitution parameters and demand parameters for each good, which results in a mixed random utility model similar to McFadden and Train (2000) and Berry, Levinsohn and Pakes (1995).

In particular, we partition consumers into different types indexed by $r \in \{1, \ldots, R\}$. The utility of an individual i of type r who consumes C_{ik}^{r} units of product k is:

$$\mathbb{U}_{i}^{r} = z_{ik}^{r} \varphi_{k}^{r} C_{ik}^{r}, \tag{31}$$

where φ_{k}^{r} captures type-r consumers' common tastes for product k; z_{ik}^{r} captures idiosyncratic consumer tastes for each product; and we have omitted the time subscript t on each variable to simplify notation. Each consumer i of type r, therefore, chooses C_{ik}^{r} units of good k to maximize utility. Since the consumer only consumes their preferred good, their budget constraint implies that $C_{ik}^{r} = E_{i}^{r} / P_{k}^{r}$, where E_{i}^{r} is the consumer's expenditure, and P_{k}^{r} is the price of the good, where we allow different types to potentially face different prices. Using this result, utility (31) can be re-written in the indirect form as:

$$\mathbb{U}_{i}^{r} = z_{ik}^{r} \left(\varphi_{k}^{r} / P_{k}^{r} \right) E_{i}^{r}. \tag{32}$$

These idiosyncratic tastes are assumed to have a Fréchet (Type-II Extreme Value) distribution:

$$G\left(z\right) = e^{-z^{-\theta^{r}}}, \tag{33}$$

where we allow the shape parameter determining the dispersion of idiosyncratic tastes (θ^{r}) to vary across types. We normalize the scale parameter of the Fréchet distribution to one, because it affects consumer expenditure shares isomorphically to the consumer tastes parameter φ_k^r.[13] Using the monotonic relationship between idiosyncratic tastes and utility in equation (32), the probability that an individual i of type r chooses product k is the same across all individuals of that type and equal to:

$$S_{ik}^r = S_k^r = \frac{\left(P_k^r/\varphi_k^r\right)^{-\theta^r}}{\sum_{\ell=1}^N \left(P_\ell^r/\varphi_\ell^r\right)^{-\theta^r}}, \tag{34}$$

as shown in Section A.7 of the web appendix. This probability (34) also equals the share of product k in the expenditure of consumers of type r (S_k^r), since all consumers of the same type are assumed to have the same expenditure: $E_i^r = E^r$. The expected utility of consumer i of type r is:

$$\mathbb{E}\left[\mathbb{U}^r\right] = \gamma^r \left[\sum_{k=1}^N \left(E_i^r\right)^{\theta^r} \left(P_k^r/\varphi_k^r\right)^{-\theta^r}\right]^{\frac{1}{\theta^r}}, \qquad \gamma^r = \Gamma\left(\frac{\theta^r - 1}{\theta^r}\right), \tag{35}$$

where $\Gamma\left(\cdot\right)$ is the Gamma function, as also shown in Section A.7 of the web appendix. This expected utility can be re-written as:

$$\mathbb{E}\left[\mathbb{U}^r\right] = \frac{E_i^r}{\mathbb{P}^r}, \tag{36}$$

where $\mathbb{P}^r$ is the unit expenditure function for consumers of type r:

$$\mathbb{P}^r = \left(\gamma^r\right)^{-1} \left[\sum_{k=1}^N \left(P_k^r/\varphi_k^r\right)^{-\theta^r}\right]^{-\frac{1}{\theta^r}}. \tag{37}$$

Note that if we change notation and define $\theta^r = \sigma^r - 1$ and assume that there is only one type (r) of consumers, equations (34) and (37) become identical to the demand system and unit expenditure function that we derived in the CES case (up to a normalization or choice of units in which to measure φ_{kt}^r to absorb the constant $\left(\gamma^r\right)^{-1}$). Therefore, the CES demand system and its "love-of-variety" property can be thought of as a means of aggregating "ideal-type" consumers who only consume one of each type of variety. We thus have a generalization of our framework to accommodate multiple types of consumers with different substitution parameters and demand parameters for each good.

Proposition 6. *Given data on prices and expenditure of consumers of each type r, the mixed random utility model defined by the indirect utility function (31) and Type-II Extreme Value distributed idiosyncratic tastes (33) with shape parameter θ^r is isomorphic to a constant elasticity of substitution (CES) model in which consumers of different types r have different demand parameters (φ_k^r) and elasticities of substitution (σ^r). This mixed random utility model implies a demand system (34) and unit expenditure function (37) for consumers of a given type r that are isomorphic (up to a normalization or choice of units for φ_k^r) to those in a mixed CES model with multiple consumer types, where $\theta^r = \sigma^r - 1$.*

Proof. The proposition follows immediately from the demand system (34) and unit expenditure function (37), substituting $\theta^r = \sigma^r - 1$. □

[13] Although we assume a Fréchet (Type-II Extreme Value) distribution for idiosyncratic tastes, because the derivations are more direct, analogous results hold in a closely-related specification with a logit (Type-I Extreme Value) distribution.

In this specification with multiple types of consumers, our unified price index now provides the exact price index for each type of consumers that allows for the entry and exit of goods over time, changes in demand for each good over time (where demand for each good and time period can now differ across consumer types) and imperfect substitutability between goods (where the degree of substitutability between goods can now vary across consumer types).

Proposition 7. *The "unified price index" (UPI) for consumer type r—which is exact for the mixed random utility model defined by the indirect utility function (31) and Type-II Extreme Value distributed idiosyncratic tastes (33)—is given by*

$$\Phi^{Ur}_{t-1,t} = \underbrace{\left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\theta^r}}}_{\text{Variety Adjustment}} \underbrace{\left[\frac{\tilde{P}^*_t}{\tilde{P}^*_{t-1}}\left(\frac{\tilde{S}^*_t}{\tilde{S}^*_{t-1}}\right)^{\frac{1}{\theta^r}}\right]}_{\text{Common-Goods Price Index } \Phi^{CGr}_{t-t,t}} . \tag{38}$$

Proof. The proposition follows from combining the expenditure share (34) and unit expenditure function (37) for each consumer type r, following the same line of argument as for the CES specification with a representative consumer in Section 3.2. □

In summary, our price index has the same functional form but a slightly different interpretation in a random utility model. While it is not valid for any individual consumer, who has idiosyncratic tastes, our index tells us the average movement in the unit expenditure function for consumers of any given type.

6 Estimation of the Elasticity of Substitution

We now turn to the estimation of the elasticity of substitution (σ). We first review existing approaches based on demand systems estimation in Section 6.1. We next develop our "reverse-weighting" (RW) estimator and characterize its asymptotic properties in Section 6.2. Finally, we provide Monte Carlo evidence on our estimator's finite sample performance in Section 6.3. One of the key results of this section is that the RW estimator can be used to derive upper and lower bounds for the true elasticity of substitution regardless of the correlation between demand and supply shocks.

6.1 Demand Systems Estimation

Taking logarithms in the common goods expenditure share (13), and first differencing over time, we obtain the following regression relationship between log changes in expenditure shares and prices for a pair of periods t and $t-1$:

$$\Delta^T \ln S^*_{kt} = \beta_0 + \beta_1 \Delta^T \ln P_{kt} + u_{kt}, \tag{39}$$

where $\beta_0 = (\sigma - 1)\,\Delta^T \ln \mathbb{P}^*_t$; $u_{kt} = (\sigma - 1)\,\Delta^T \ln \varphi_{kt}$; and Δ^T is the time-difference operator such that $\Delta^T \ln P_{kt} = \ln (P_{kt}/P_{kt-1})$.

The main challenge in estimating the demand system (39) is that shocks to prices ($\Delta^T \ln P_{kt}$) can be correlated with shocks to demand ($u_{kt} = (\sigma - 1)\,\Delta^T \ln \varphi_{kt}$), in which case the OLS estimate of the parameter β_1 is inconsistent and subject to omitted variable bias: $\hat{\beta}_1^{OLS} = (1-\sigma)\left(1 - \text{cov}\left(\Delta^T \ln P_{kt}, \Delta^T \ln \varphi_{kt}\right)/\text{var}\left(\Delta^T \ln P_{kt}\right)\right)$. The standard approach to this problem is to specify a supply-side such as:

$$\Delta^T \ln P_{kt} = \gamma_0 + \gamma_1 \Delta^T \ln S^*_{kt} + \gamma_2 Z_{kt} + v_{kt}, \tag{40}$$

and to search for instruments (Z_{kt}) that are both strongly correlated with the log change in prices (cov $\left(\Delta^T \ln P_{kt}, Z_{kt}\right) \neq 0$) and have no direct effect on expenditure shares (cov $(u_{kt}, Z_{kt}) = 0$).

The main alternative approach is that of Feenstra (1994). This alternative estimator second differences the demand system (39) and supply-system (40) across goods ($\Delta^{T,K} \ln P_{kt} = \ln\left(P_{kt}/P_{kt-1}\right) - \ln\left(P_{\ell t}/P_{\ell t-1}\right)$), and uses the identifying assumption that the double-differenced demand and supply shocks are orthogonal to one another and heteroskedastic. The assumption that the shocks are orthogonal defines a rectangular hyperbola in the demand-supply elasticity space for each good. The assumption of heteroskedasticity across goods implies that these rectangular hyperbolas for different goods do not lie on top of another. Therefore, their intersection separately identifies the demand and supply elasticities.

Both these estimators solve an underidentification problem in the demand system for expenditure shares in equation (39). We have $N_{t,t-1} - 1$ independent equations for the change in log expenditure shares ($\Delta^T \ln S^*_{kt}$) in this system. But there are $N_{t,t-1} - 1$ independent demand shocks ($\Delta^T \ln \varphi_{kt}$), up to a normalization, plus one elasticity of substitution (σ) to be estimated. The instrumental variables estimator achieves identification by adding additional equations for the orthogonality of the demand shocks and the instruments. Similarly, the Feenstra (1994) estimator achieves identification by incorporating additional equations for the orthogonality and heteroskedasticity of the double-differenced demand shocks and supply shocks for each good. When the identifying assumptions of each of these approaches are satisfied, they yield consistent estimates of the elasticity of substitution (σ) that can be used in our unified price index.

6.2 The Reverse-Weighting Estimator

However, there can be settings in international trade and macroeconomics, in which a researcher is concerned about making the assumption that double-differenced demand and supply shocks are orthogonal, but does not have access to the detailed information required to construct valid instruments for every single sector in the aggregate economy. Nonetheless, the researcher may need to construct price aggregates that depend on demand-side parameters. We now develop a new estimator of the elasticity of substitution that is appropriate for such settings, which combines the demand system with the unit expenditure function. We show that the functional form of CES preferences can be used to provide upper and lower bounds to the true elasticity of substitution regardless of the correlation of demand and supply shocks. This estimator requires only price and expenditure share data and is easy to implement across a broad range of sectors for which it may be infeasible to construct valid instrumental variables. Although our estimator yields set rather than point identification, we show that these bounds are tight in both our Monte Carlos and the observed data.

We begin by rewriting our forward and backward differences of the CES price index in terms of aggregate demand shifters that summarize the effect of changes in demand for each good on aggregate utility. Using these forward and backward differences (equations (20) and (21) respectively), the common good expenditure share (13), and the unified price index (16), we obtain the following system of three equivalent expressions for the change in the cost of living from period $t-1$ to t:

$$\frac{\mathbb{P}_t}{\mathbb{P}_{t-1}} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}} \Theta^F_{t-1,t} \left[\sum_{k\in\Omega_{t,t-1}} S^*_{kt-1}\left(\frac{P_{kt}}{P_{kt-1}}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}, \tag{41}$$

$$\frac{\mathbb{P}_t}{\mathbb{P}_{t-1}} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}} \left(\Theta^B_{t,t-1}\right)^{-1} \left[\sum_{k\in\Omega_{t,t-1}} S^*_{kt}\left(\frac{P_{kt}}{P_{kt-1}}\right)^{-(1-\sigma)}\right]^{-\frac{1}{1-\sigma}}, \tag{42}$$

$$\frac{\mathbb{P}_t}{\mathbb{P}_{t-1}} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}\right)^{\frac{1}{\sigma-1}} \frac{\tilde{P}^*_t}{\tilde{P}^*_{t-1}} \left(\frac{\tilde{S}^*_t}{\tilde{S}^*_{t-1}}\right)^{\frac{1}{\sigma-1}}, \tag{43}$$

where the forward and backward aggregate demand shifters ($\Theta^F_{t-1,t}$ and $\Theta^B_{t,t-1}$) can be written as:

$$\Theta^F_{t-1,t} \equiv \left[\frac{\sum_{k\in\Omega_{t,t-1}} S^*_{kt-1}\left(\frac{P_{kt}}{P_{kt-1}}\right)^{1-\sigma}\left(\frac{\varphi_{kt}}{\varphi_{kt-1}}\right)^{\sigma-1}}{\sum_{k\in\Omega_{t,t-1}} S^*_{kt-1}\left(\frac{P_{kt}}{P_{kt-1}}\right)^{1-\sigma}}\right]^{\frac{1}{1-\sigma}} = \left[\sum_{k\in\Omega_{t,t-1}} S^*_{kt}\left(\frac{\varphi_{kt-1}}{\varphi_{kt}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}}, \tag{44}$$

$$\Theta^B_{t,t-1} \equiv \left[\frac{\sum_{k\in\Omega_{t,t-1}} S^*_{kt}\left(\frac{P_{kt-1}}{P_{kt}}\right)^{1-\sigma}\left(\frac{\varphi_{kt-1}}{\varphi_{kt}}\right)^{\sigma-1}}{\sum_{k\in\Omega_{t,t-1}} S^*_{kt}\left(\frac{P_{kt-1}}{P_{kt}}\right)^{1-\sigma}}\right]^{\frac{1}{1-\sigma}} = \left[\sum_{k\in\Omega_{t,t-1}} S^*_{kt-1}\left(\frac{\varphi_{kt}}{\varphi_{kt-1}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}},$$

as shown in Section A.8 of the web appendix.

The forward and backward aggregate demand shifters in (44) have an intuitive interpretation. Each aggregate demand shifter is an expenditure-share weighted average of the changes in demand for each good, where the expenditure-share weights are either the initial or the final expenditure shares. While all three formulations of the price index are equivalent, only the UPI in equation (43) is money metric, because it alone contains no demand parameters. Each of these aggregate demand shifters are functions of the elasticity of substitution, which raises the question of whether there exists an elasticity of substitution that renders all three of our formulas for the CES price index consistent with the same money-metric utility function. The reverse-weighting estimator imposes this identifying assumption, which corresponds to:

$$\Theta^F_{t-1,t} = \left(\Theta^B_{t,t-1}\right)^{-1} = 1. \tag{45}$$

We use this identifying assumption to construct a generalized method of moments (GMM) estimator of the elasticity of substitution (σ). Combining the equalities (41)-(43), we obtain the following moment functions:

$$M(\sigma) \equiv \begin{pmatrix} m^F_t(\sigma) \\ m^B_t(\sigma) \end{pmatrix} \equiv \begin{pmatrix} \frac{1}{1-\sigma}\ln\left[\sum_{k\in\Omega_{t,t-1}} S^*_{kt-1}\left(\frac{P_{kt}}{P_{kt-1}}\right)^{1-\sigma}\right] - \ln\left[\frac{\tilde{P}^*_t}{\tilde{P}^*_{t-1}}\left(\frac{\tilde{S}^*_t}{\tilde{S}^*_{t-1}}\right)^{\frac{1}{\sigma-1}}\right] \\ -\frac{1}{1-\sigma}\ln\left[\sum_{k\in\Omega_{t,t-1}} S^*_{kt}\left(\frac{P_{kt}}{P_{kt-1}}\right)^{-(1-\sigma)}\right] - \ln\left[\frac{\tilde{P}^*_t}{\tilde{P}^*_{t-1}}\left(\frac{\tilde{S}^*_t}{\tilde{S}^*_{t-1}}\right)^{\frac{1}{\sigma-1}}\right] \end{pmatrix} = \begin{pmatrix} -\ln\left(\Theta^F_{t-1,t}\right) \\ \ln\left(\Theta^B_{t,t-1}\right) \end{pmatrix}. \tag{46}$$

The reverse weighting estimator ($\hat{\sigma}^{RW}$) solves:

$$\hat{\sigma}^{RW} = \arg\min\left\{M\left(\sigma^{RW}\right)' \times \mathbb{I} \times M\left(\sigma^{RW}\right)\right\}, \tag{47}$$

where we weight the two moments for the forward and backward difference equally by using the identity matrix ($\mathbb{I}$) for the weighting matrix.[14]

We term this estimator the "reverse-weighting" (RW) estimator ($\hat{\sigma}^{RW}$), because it involves equating expressions for the change in the cost of living using both initial-period and final-period expenditure share weights. Our use of the identity matrix as the weighting matrix ensures that this estimator minimizes the sum of squared deviations of the aggregate demand shifters ($\left(-\ln\left(\Theta^F_{t-1,t}\right)\right)^2 + \left(\ln\left(\Theta^B_{t,t-1}\right)\right)^2$) from zero. Therefore it chooses the elasticity of substitution to minimize the squared deviations from money-metric utility for the forward and backward differences of the price index.

The reverse-weighting estimator is overidentified, because we have two moment conditions ($m^F_t(\sigma) = 0$ and $m^B_t(\sigma) = 0$ in equation (46)) to identify one parameter (σ). Therefore, we can also examine an exactly-identified "forward estimator" ($\hat{\sigma}^F$), in which we use only the forward moment condition ($m^F_t(\sigma) = 0$), and an exactly-identified "backward estimator" ($\hat{\sigma}^B$), in which we use only the backward moment condition ($m^B_t(\sigma) = 0$). Given our reverse-weighting estimate of the elasticity of substitution ($\hat{\sigma}^{RW}$), the demand system in equation (15) provides a system of $N_{t,t-1} - 1$ equations in the $N_{t,t-1} - 1$ independent demand shifters (up to our normalization). Therefore, we can use the demand system to uniquely determine the demand parameter for each good k and period t ($\hat{\varphi}^{RW}_{kt}$) up to our common choice of units in which to measure the demand parameters ($\tilde{\varphi}_t = \bar{\varphi}$).

We now show that the RW estimator ($\hat{\sigma}^{RW}$) has an attractive economic interpretation. First, the money-metric forward difference (equation (41) with $\Theta^F_{t-1,t} = 1$) corresponds to the change in the cost of living evaluated using period $t-1$ tastes. Second, the money-metric backward difference (equation (42) with $\Theta^B_{t,t-1} = 1$) corresponds to the change in the cost of living evaluated using period t tastes. Therefore, the RW estimator minimizes the sum of squared deviations between (i) the unified price index evaluated using tastes in each period (inverting the demand system), (ii) the change in the cost of living evaluated using period $t-1$ tastes, and (iii) the change in the cost of living using period t tastes.

Proposition 8. *The reverse-weighting estimator ($\hat{\sigma}^{RW}$) minimizes the sum of squared deviations between the money-metric unified price index (UPI) evaluated using tastes in each period, the change in the cost of living evaluated using period $t-1$ tastes, and the change in the cost of living using period t tastes.*

Proof. See Section A.10 of the web appendix. □

Although this property of the RW estimator is attractive for making welfare comparisons over time, a natural question to ask is the conditions under which it consistently estimates the true elasticity of substitution. Therefore, we now provide analytical results on the asymptotic properties of the RW estimator. The RW estimator exploits variation in expenditure-share-weighted changes in prices, which determine the slopes of the three equivalent expressions for the price index in equations (41)-(43) with respect to σ. Therefore, we require $\sigma \neq 1$, in order for expenditure shares to vary with relative prices. Additionally, we exclude the knife-edge case in which all common goods have the same expenditure share in both time periods:

[14] In Section A.9 of the web appendix, we show that the reverse-weighting estimator in equations (46) and (47) generalizes to allow for a Hicks-neutral shifter of tastes that is common to all goods because, like the variety correction term, this Hicks-neutral shifter cancels from the three equivalent expressions for the change in the cost of living in equations (41)-(43).

$S^*_{kt-1} = S^*_{kt} = 1/N_{t,t-1}$. Outside of this knife-edge case, the UPI and the forward and backward differences of the price index in equations (41)-(43) have different slopes with respect to σ. Hence, they satisfy a single-crossing property, which identifies the elasticity of substitution, as shown formally in the proof of the propositions below.

First, we show that the RW estimator consistently estimates the true elasticity of substitution as the demand shocks for each good become small.

Proposition 9. *As changes in demand become small ($(\varphi_{kt}/\varphi_{kt-1}) \to 1$), the reverse-weighting estimator consistently estimates the true elasticity of substitution ($\hat{\sigma}^{RW} \xrightarrow{p} \sigma$) and demand ($\hat{\varphi}^{RW}_{kt} \xrightarrow{p} \varphi_{kt}$) for each good k in each time period t.*

Proof. See Section A.11 of the web appendix. □

As the demand shocks for each good become small ($(\varphi_{kt}/\varphi_{kt-1}) \to 1$), the forward and backward aggregate demand shifters in equation (44) converge to one ($\Theta^F_{t-1,t} \xrightarrow{p} 1$ and $\Theta^B_{t,t-1} \xrightarrow{p} 1$), and the assumption that all three expression for the price index are consistent with the same money-metric utility function is satisfied ($\Theta^F_{t-1,t} = \left(\Theta^B_{t,t-1}\right)^{-1} = 1$). In this case, the forward and backward differences of the price index reduce to the expenditure-share-weighted average of the price changes, and hence take a money-metric form. Therefore the RW estimator consistently estimates the elasticity of substitution (σ) and demand parameter (φ_{kt}) for each good. More generally, the assumption of money-metric utility is satisfied up to a first-order approximation, as shown in section A.12 of the web appendix. Hence the RW estimator can be interpreted as providing a first-order approximation to the data.

Second, we show that the RW estimator consistently estimates the elasticity of substitution as the number of common goods becomes large ($N_{t,t-1} \to \infty$) if demand shocks are independently and identically distributed across goods.

Proposition 10. *Assuming that demand shocks are independently and identically distributed, $(\varphi_{kt}/\varphi_{kt-1}) \sim$ i.i.d. $\left(1, \chi^2_{\varphi}\right)$ for $(\varphi_{kt}/\varphi_{kt-1}) \in (0, \infty)$, as the number of common goods becomes large ($N_{t,t-1} \to \infty$), the reverse-weighting estimator consistently estimates the elasticity of substitution ($\hat{\sigma}^{RW} \xrightarrow{p} \sigma$) and the demand parameter for each good k and period t ($\hat{\varphi}^{RW}_{kt} \xrightarrow{p} \varphi_{kt}$).*

Proof. See Section A.13 of the web appendix. □

When demand shocks are independently and identically distributed and the number of common goods is large, demand shocks average out across all common goods. Therefore, these demand shocks for individual goods have no direct effect on the forward and backward differences of the price index. In this case, our identifying assumption that the forward and backward differences of the price index are money metric ($\Theta^F_{t-1,t} = \left(\Theta^B_{t,t-1}\right)^{-1} = 1$) is again satisfied. Hence the RW estimator consistently estimates the elasticity of substitution (σ) and the demand parameter for each good (φ_{kt}).

Third, when demand and price shocks are correlated ($\text{cov}\left((\varphi_{kt}/\varphi_{kt-1}), (P_{kt}/P_{kt-1})\right) \neq 0$), we show that the RW estimator can be used to derive upper and lower bounds to the true elasticity of substitution. To determine these bounds, we begin by characterizing the patterns of departures from money metric utility as a function of the correlation between demand and price shocks.

Proposition 11. *As the number of common goods becomes large ($N_{t,t-1} \to \infty$), a positive correlation between demand and price shocks ($cov\left(\left(\varphi_{kt}/\varphi_{kt-1}\right), \left(P_{kt}/P_{kt-1}\right)\right) > 0$) implies that the forward aggregate- demand shifter is strictly greater than one and the backward aggregate-demand shifter is strictly less than one: $\Theta^F_{t-1,t} > 1 > \Theta^B_{t,t-1}$. A negative correlation ($cov\left(\left(\varphi_{kt}/\varphi_{kt-1}\right), \left(P_{kt}/P_{kt-1}\right)\right) < 0$) implies that the converse is true: $\Theta^F_{t-1,t} < 1 < \Theta^B_{t,t-1}$.*

Proof. See Section A.14 of the web appendix. □

When demand and price shocks are positively correlated, consumers on average demand more of the goods whose price has risen. This magnifies the negative impact of these price increases on the cost of living, because these goods account for a larger share of expenditure than they would in the absence of the demand increase. Therefore, the true increase in the cost of living from period $t-1$ to t (incorporating both demand and price changes) is strictly greater than it would be if tastes were held constant at their initial values ($\Theta^F_{t-1,t} > 1$). Similarly, the true reduction in the cost of living from period t to $t-1$ (incorporating both demand and price changes) is smaller than it would be if tastes were held constant at their final values ($\Theta^B_{t,t-1} < 1$). The reason is analogous: although the goods whose price increased over time were cheaper in period $t-1$ than in period t, they were also demanded less by the consumer in period $t-1$ than in period t.

In contrast, when demand and price shocks are negatively correlated, consumers demand less of the goods whose price is risen. This mitigates the negative impact of these price increases on the cost of living, because these goods account for a smaller share of expenditure than they would in the absence of the demand increase. Hence, the true increase in the cost of living from period $t-1$ to t (taking account of both demand and price changes) is strictly lower than it would be if tastes were held constant at their initial values ($\Theta^F_{t-1,t} < 1$). Similarly, the true reduction in the cost of living from period t to $t-1$ (incorporating both demand and price changes) is larger than it would be if tastes were held constant at their final values ($\Theta^B_{t,t-1} > 1$). The reasoning takes the same form: the goods whose price increased over time were both cheaper and demanded more by the consumer in period $t-1$ than in period t.

We now characterize the impact of these departures from money-metric utility on the RW estimate of the elasticity of substitution.

Proposition 12. *As the number of common goods becomes large ($N_{t,t-1} \to \infty$), a positive correlation between demand and price shocks ($cov\left(\varphi_{kt}/\varphi_{kt-1}, P_{kt}/P_{kt-1}\right) > 0$) implies that the reverse-weighting estimate of the elasticity of substitution is strictly less than its true value ($\hat{\sigma}^{RW} < \sigma$), while a negative correlation between demand and price shocks ($cov\left(\varphi_{kt}/\varphi_{kt-1}, P_{kt}/P_{kt-1}\right) < 0$) implies that the converse is true ($\hat{\sigma}^{RW} > \sigma$).*

Proof. See section A.15 of the web appendix. □

As shown in Proposition 11, when demand and price shocks are positively correlated, the increase in the cost of living incorporating changes in tastes is greater than that evaluated using initial tastes ($\Theta^F_{t-1,t} > 1$)

and smaller than that using final period tastes ($\Theta^{B}_{t,t-1} < 1$). A larger value of the elasticity of substitution dampens the impact of changes in tastes in raising the forward aggregate-demand shifter above one and reducing the backward aggregate-demand shifter below one. Hence, the RW estimator compensates for the correlation between demand and price shocks by increasing the elasticity of substitution, so as to reduce the forward aggregate-demand shifter ($\Theta^{F}_{t-1,t}$) and increase the backward aggregate-demand shifter ($\Theta^{B}_{t,t-1}$), taking them as close to one as possible. In contrast, when demand and price shocks are negatively correlated, the increase in the cost of living incorporating changes in tastes is smaller than that evaluated using initial tastes ($\Theta^{F}_{t-1,t} < 1$) and greater than that using final period tastes ($\Theta^{B}_{t,t-1} > 1$). To compensate, the RW estimator reduces the elasticity of substitution, so as to increase the forward aggregate demand shifter ($\Theta^{F}_{t-1,t}$) and reduce the backward aggregate demand shifter ($\Theta^{B}_{t,t-1}$), taking them as close to one as possible.

Together, Propositions 11 and 12 imply that the RW estimate ($\hat{\sigma}^{RW}$) is below the true value of the elasticity of substitution when there is a positive correlation between demand and price shocks and above this true value for a negative correlation. To provide upper and lower bounds for the true elasticity of substitution, we now construct an estimator that has the opposite pattern of departures from the true parameter value, following an approach used in another context in the literature on reverse regressions (e.g., Liviatan 1961 and Klepper and Leamer 1984). From the expressions for the aggregate demand shifters in equation (44), the RW moment conditions in equations (45) and (46) can be equivalently written as:

$$\Theta^{F}_{t-1,t} - 1 = \left[\frac{\sum_{k\in\Omega_{t,t-1}} S^{*}_{kt-1} \left(\frac{P_{kt}}{P_{kt-1}}\right)^{1-\sigma} \left(\frac{\varphi_{kt}}{\varphi_{kt-1}}\right)^{\sigma-1}}{\sum_{k\in\Omega_{t,t-1}} S^{*}_{kt-1} \left(\frac{P_{kt}}{P_{kt-1}}\right)^{1-\sigma}} \right]^{\frac{1}{1-\sigma}} - 1 = 0, \tag{48}$$

$$\left(\Theta^{B}_{t,t-1}\right)^{-1} - 1 = \left[\frac{\sum_{k\in\Omega_{t,t-1}} S^{*}_{kt} \left(\frac{P_{kt-1}}{P_{kt}}\right)^{1-\sigma} \left(\frac{\varphi_{kt-1}}{\varphi_{kt}}\right)^{\sigma-1}}{\sum_{k\in\Omega_{t,t-1}} S^{*}_{kt} \left(\frac{P_{kt-1}}{P_{kt}}\right)^{1-\sigma}} \right]^{-\frac{1}{1-\sigma}} - 1 = 0,$$

where we can use our inversion of the demand system in equation (15) to substitute for the demand shocks ($\varphi_{kt}/\varphi_{kt-1}$) in terms of observed data and parameters.

Using this alternative formulation of the RW moment conditions, we construct a "double-reverse weighting" (DRW) estimator that uses the inverse of the demand shocks ($(\varphi_{kt}/\varphi_{kt-1})^{-1}$) instead of the actual demand shocks ($\varphi_{kt}/\varphi_{kt-1}$):

$$\left[\frac{\sum_{k\in\Omega_{t,t-1}} S^{*}_{kt-1} \left(\frac{P_{kt}}{P_{kt-1}}\right)^{1-\sigma} \left(\frac{\varphi_{kt}}{\varphi_{kt-1}}\right)^{-(\sigma-1)}}{\sum_{k\in\Omega_{t,t-1}} S^{*}_{kt-1} \left(\frac{P_{kt}}{P_{kt-1}}\right)^{1-\sigma}} \right]^{\frac{1}{1-\sigma}} - 1 = 0, \tag{49}$$

$$\left[\frac{\sum_{k\in\Omega_{t,t-1}} S^{*}_{kt} \left(\frac{P_{kt-1}}{P_{kt}}\right)^{1-\sigma} \left(\frac{\varphi_{kt-1}}{\varphi_{kt}}\right)^{-(\sigma-1)}}{\sum_{k\in\Omega_{t,t-1}} S^{*}_{kt} \left(\frac{P_{kt-1}}{P_{kt}}\right)^{1-\sigma}} \right]^{-\frac{1}{1-\sigma}} - 1 = 0,$$

where we can again use our inversion of the demand system in equation (15) to substitute for the inverse demand shocks ($(\varphi_{kt}/\varphi_{kt-1})^{-1}$) in terms of observed data and parameters.

Equations (48) and (49) are identical except for the negative sign in the exponents on the demand shocks ($\varphi_{kt}/\varphi_{kt-1}$) in equation (49). Whenever the actual demand shocks ($\varphi_{kt}/\varphi_{kt-1}$) take large values above one, the inverse demand shocks ($(\varphi_{kt}/\varphi_{kt-1})^{-1}$) necessarily take small values below one. Therefore, the correlation between the inverse demand shocks ($(\varphi_{kt}/\varphi_{kt-1})^{-1}$) and price shocks ($P_{kt}/P_{kt-1}$) in equation (49) has the opposite sign from that between the actual demand shocks ($\varphi_{kt}/\varphi_{kt-1}$) and price shocks ($P_{kt}/P_{kt-1}$) in equation (48). Since it is these correlations that generate the departures between the RW estimator and the true parameter value, it follows that the DRW estimator has the opposite pattern of departures from the true parameter value to the RW estimator. In other words, if the RW estimator underestimates the true elasticity, the DRW estimator will overestimate it and *vice versa*.

We now use this property to generate bounds for the true elasticity. From Propositions 11 and 12, whenever demand and price shocks are positively correlated, the RW estimator is biased downward, which implies that the DRW estimator is biased upward. Similarly, whenever demand and price shocks are negatively correlated, the RW estimator is biased upward, which implies that the DRW estimator is biased downward. A first important implication of these results is that the relative value of the two estimators reveals the pattern of correlation between demand and price shocks. If the two estimators take the same value, demand and price shocks are orthogonal; if the RW estimator lies above the DRW estimator, demand and price shocks are negatively correlated; and if the DRW estimator lies above the RW estimator, demand and price shocks are positively correlated. A second important implication is that the maximum of the two estimators ($\max\left\{\hat{\sigma}^{RW}, \hat{\sigma}^{DRW}\right\}$) provides an upper bound to the true elasticity of substitution, while the minimum of them ($\min\left\{\hat{\sigma}^{RW}, \hat{\sigma}^{DRW}\right\}$) provides a lower bound to the true elasticity of substitution.

6.3 Monte Carlo

We now provide Monte Carlo evidence on the finite sample performance of the RW estimator. We assume a model economy with CES demand and a conventional supply side in the form of monopolistic competition and constant marginal costs (as in Krugman 1980 and Melitz 2003). We first assume true values for the model's parameters (the elasticity of substitution σ) and its structural residuals (demand and marginal cost for each good). We next solve for equilibrium prices and expenditure shares in this economy. Finally, we suppose that a researcher only observes data on these prices and expenditure shares and implements our unified approach. For each combination of parameters, we undertake 250 replications of the model. We compare the mean and standard deviation of the parameter estimates across these replications with the true parameter values.

As the RW estimator uses only the subset of common goods, we focus on this subset, and are not required to make assumptions about entering and exiting goods. We consider numbers of common goods ranging from 10 to 1,000. We focus for simplicity on a single pair of time periods $t-1$ and t. We assume the following values for the model's parameters. We set the elasticity of substitution equal to 4, which is consistent with estimates using U.S. data in Bernard, Eaton, Jensen and Kortum (2004). The time-varying demand shifters (φ_{kt}) and marginal cost shifters (b_{kt}) are drawn from a joint log normal distribution:

$$\begin{pmatrix} \ln \varphi_{kt} \\ \ln b_{kt} \end{pmatrix} \sim N\begin{pmatrix} 0 & \chi_{\varphi}^2 & \rho_{\varphi P}\chi_{\varphi}\chi_P \\ 0 & \rho_{\varphi P}\chi_{\varphi}\chi_P & \chi_P^2 \end{pmatrix}.$$

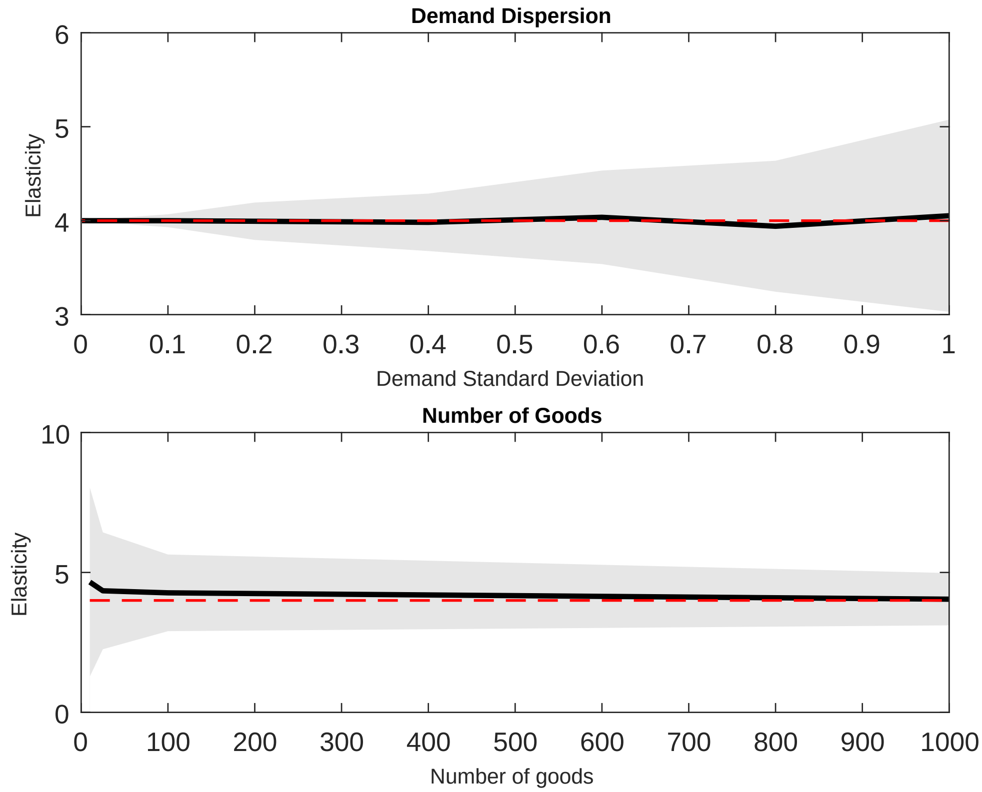
We allow demand ($\ln \varphi_{kt}$) and marginal cost ($\ln b_{kt}$) to be correlated across goods (when $\rho_{\varphi P} \neq 0$), but assume that this joint log-normal distribution is independent across time periods. Note that the difference between two normal distributions is normally distributed. Therefore, both demand shocks ($\ln\left(\varphi_{kt}/\varphi_{kt-1}\right)$) and marginal cost shocks ($\ln\left(b_{kt}/b_{kt-1}\right)$) are joint log normally distributed, and can be correlated across goods (when $\rho_{\varphi P} \neq 0$). Finally, with a constant elasticity of substitution and monopolistic competition, price shocks ($\ln\left(P_{kt}/P_{kt-1}\right)$) are exactly proportional to marginal cost shocks ($\ln\left(b_{kt}/b_{kt-1}\right)$).[15]

We begin by considering the case in which demand and marginal costs are independently distributed across goods ($\rho_{\varphi P} = 0$). In our first Monte Carlo exercise, we vary the dispersion of demand shocks (as in Proposition 9 above), holding constant the dispersion of marginal cost shocks and the number of common goods. We assume 1,000 common goods, a standard deviation of log marginal costs of 1, and a standard deviation of log demand ranging from 0.001 to 1. In the top panel of Figure 2, we show the mean of the RW estimate ($\hat{\sigma}^{RW}$) across the 250 replications (solid black line), the 95 percent confidence intervals (gray shading), and the true parameter value (red dashed line). We find that the mean RW estimate is always close to the true parameter value and we are unable to reject the null hypothesis that it is equal to the true parameter for all values of the dispersion of demand. As we reduce the dispersion of demand, the standard deviation of the RW estimate across replications falls, as reflected in the narrowing of the confidence intervals.

In our second Monte Carlo exercise, we retain the assumption of independent demand and marginal cost shocks, and vary the number of common goods (as in Proposition 10), holding constant the dispersions of demand and marginal costs. We assume standard deviations of log demand and log marginal cost of 1 and vary the number of common goods from 10 to 1,000. In the bottom panel of Figure 2, we show the mean of the RW estimate ($\hat{\sigma}^{RW}$) across the 250 replications (solid black line), the 95 percent confidence intervals (gray shading), and the true parameter value (red dashed line). We find relatively small differences between the RW estimate and the true parameter value for small numbers of common goods, such as 10 or 25, and we are we are unable to reject the null hypothesis that the RW estimate is equal to the true parameter value for all numbers of common goods. As we increase the number of common goods, the standard deviation of the RW estimate across replications again declines.

[15] We assume that firms find it profitable to supply all common goods in each time period, which implicitly corresponds to an assumption that any fixed costs of production are sufficiently small relative to the variable profits for each good.

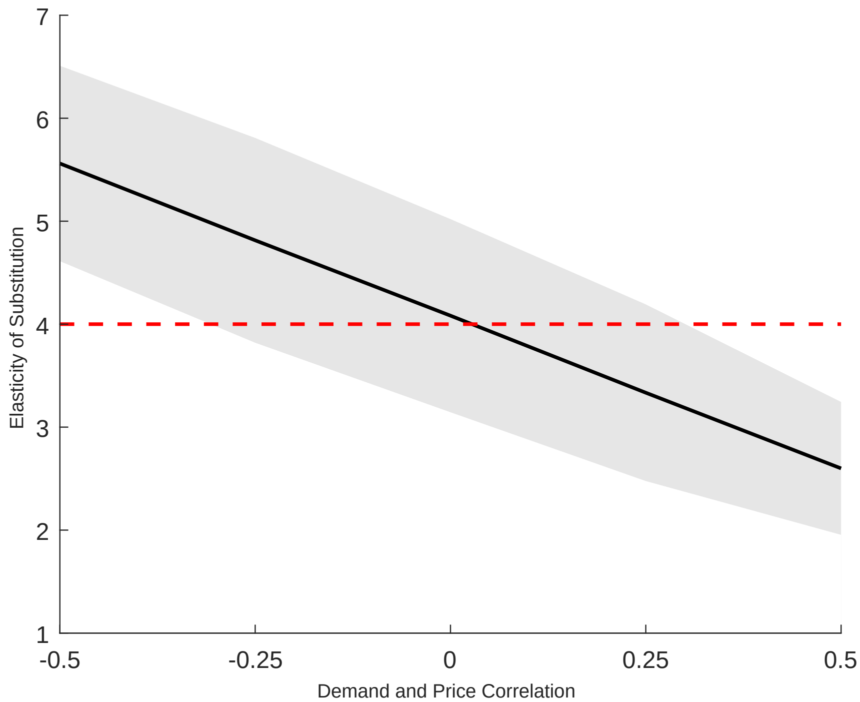
Figure 2: Reverse-weighting Estimates with Independent Demand and Price Shocks



We now consider the case in which demand and marginal costs shocks are correlated across goods ($\rho_{\varphi P} \neq 0$). In our third Monte Carlo exercise, we vary this correlation (as in Propositions 11 and 12 above), holding constant the number of common goods and the dispersions of demand and marginal cost shocks. We assume 1,000 common goods, standard deviations of log demand and log marginal costs of 1, and a correlation between demand and price shocks ranging from -0.5 to 0.5. In Figure 3, we show the mean of the RW estimate ($\hat{\sigma}^{RW}$) across the 250 replications (solid black line), the 95 percent confidence intervals (gray shading), and the true parameter value (red dashed line). Consistent with the RW estimator providing a first-order approximation to the data, we find that the mean estimate remains relatively close to the true parameter value. Hence, we are unable to reject the null hypothesis of the true parameter value for correlations as large as 0.25 in absolute value. In line with Propositions 11 and 12, we find that the RW estimate is above the true elasticity for a negative correlation between demand and price shocks and below the true elasticity for a positive correlation. As these correlations increase towards 0.50 in absolute value, these differences become statistically significant at conventional critical values.

In Figure 4, we show the upper and lower bounds for the true elasticity of substitution, which are constructed as the maximum and minimum of the RW and DRW estimators respectively (labelled "Upper" and "Lower"). We also show the mid-point between these bounds (labelled "Mid"). For positive correlations, the upper bound corresponds to the DRW estimator and the lower bound equals the RW estimator. In contrast, for negative correlations, the upper bound corresponds to the RW estimator and the lower bound equals the DRW estimator. Consistent with our earlier propositions, the true value for the elasticity lies in between these

Figure 3: Reverse-weighting Estimates with Correlated Demand and Price Shocks



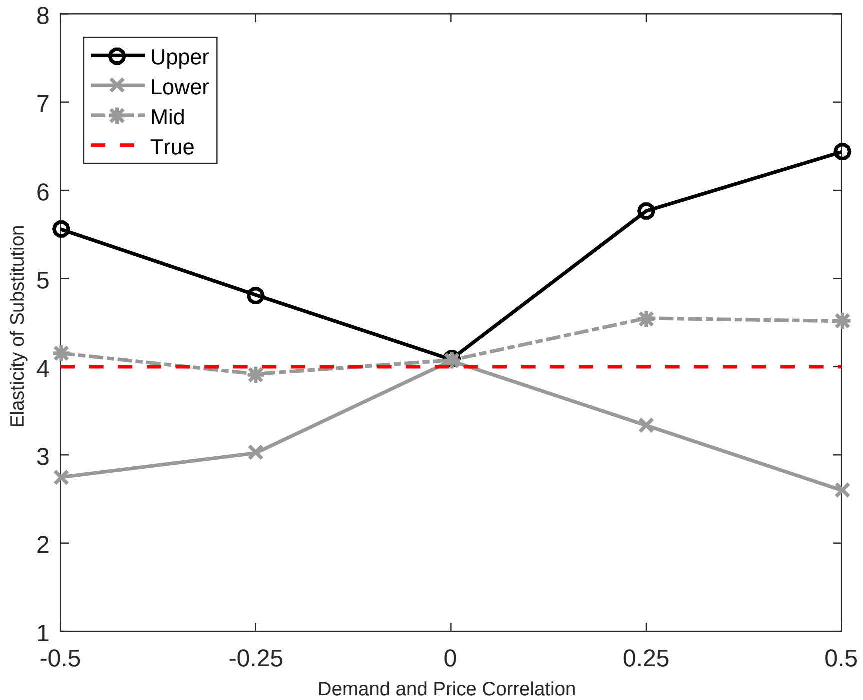
two bounds, and these bounds coincide when demand and price shocks are independently distributed. Even as the correlations become as large as 0.50 in absolute value, the bounds remain relatively tight, identifying a narrow range of possible values for the elasticity of substitution. Across correlations ranging from -0.50 to 0.50, the mid-point of these bounds remains close to the true parameter value.

Taken together, the results from these Monte Carlos are consistent with our analytical propositions, and show that our unified approach is successful in determining the true parameter value in finite samples. The RW estimator consistently estimates the true parameter when demand shocks are small and when the number of common goods is large and demand shocks are independently distributed. Together, the RW and DRW estimates provide bounds for the true elasticity of substitution. If these two estimates take the same value, demand and price shocks are orthogonal to one another. More generally, the relative value of these estimates reveals the correlation between demand and price shocks: positive correlations result in a DRW estimate above the RW estimate, and negative correlations lead to the reverse pattern. Finally, the mid-point of these bounds remains close to the true parameter value throughout all of our simulations, even for substantial correlations between demand and price shocks.

7 Results

In this section, we implement our unified price index empirically and compare the results to those using conventional price indexes. We first discuss the bar-code data used in our empirical implementation. We next estimate the elasticity of substitution for each of the product groups in our data using the reverse-weighting estimator. Finally, we compute the unified price index for each product group and in the aggregate and report

Figure 4: Upper and Lower Bounds for the True Elasticity of Substitution



the results of comparisons with existing exact and superlative price indexes (e.g., Fisher, Törnqvist) and with standard statistical price indexes (e.g., Laspeyres).

7.1 Data

We estimate the model using bar-code data from the Nielsen HomeScan database,[16] which contains price and purchase quantity data for millions of bar codes bought between 2004 and 2014. A major advantage of bar-code data over other types of price and quantity data is that product quality does not vary within a bar code, because any change in observable product characteristics results in the introduction of a new bar code. Bar codes are inexpensive to purchase and manufacturers are discouraged from reusing them because reusing the same bar code for different goods or using several bar codes for the same product can create problems for store inventory systems that inform stores about how much of each product is available. Thus, bar codes are typically unique product identifiers and changes in physical attributes manifest themselves through the creation (and destruction) of bar-coded goods, not changes in the characteristics of existing bar-coded goods. This property means that shifts in demand for bar-coded goods cannot be driven by changes in the physical quality of the good, which makes these data ideal for identifying demand shift parameters, φ_{kt}.

The data is based on a sample of approximately 50,000 households each year who scan in the price and quantity of every bar-coded good they buy each week. Nielsen adjusts the data for sampling errors (response rates that are higher or lower for different demographic groups) and enables us to compute national total

[16]Our results are calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business. Information on availability and access to the data is available at http://research.chicagobooth.edu/nielsen

value and quantity purchased of each bar-coded good. The set of goods represents close to the universe of bar-coded goods available in grocery, mass-merchandise, and drug stores, representing around a third of all goods categories included in the CPI.

Nielsen organizes goods into product groups, which are based on where goods appear in stores. We dropped "variable-weight" product groups which contain products whose quality may vary (e.g., fresh foods) and focus on the one hundred product groups constituting "packaged goods." The largest of these are carbonated beverages, pet food, paper products, bread and baked goods, and tobacco. The quantity units do not vary for bar codes and are typically defined to be volume, weight, area, length, or counts (e.g., fluid ounces for Carbonated Beverages). We also adjust for multipacks, so we compare the price per battery, not the price per package of batteries.

In choosing the time frequency with which to use the bar-code data, we face a trade-off. On the one hand, as we work with higher frequency data, we are closer to observing actual prices paid for bar-codes as opposed to averages of prices. Thus, high-frequency data has the advantage of allowing for a substantial amount of heterogeneity in price and consumption data. On the other hand, the downside is that the assumption that the total quantity purchased equals the total quantity consumed breaks down in very high-frequency data (e.g., daily or weekly) because households do not consume every item on the same day or even week they purchase it. Thus, the choice of data frequency requires a tradeoff between choosing a sufficiently high frequency that keeps us from averaging out most of the price variation, and a low enough frequency that enables us to be reasonably confident that purchase and consumption quantities are close.[17]

In order to deal with these issues, we use a quarterly frequency in our baseline specification (we find very similar results in a robustness test using an annual frequency). We therefore collapse the data to construct quarterly or annual samples of total value sold, total quantity sold, and average price. Four-quarter differences were then computed by comparing values for the fourth quarter of each year relative to the fourth quarter of the previous year.

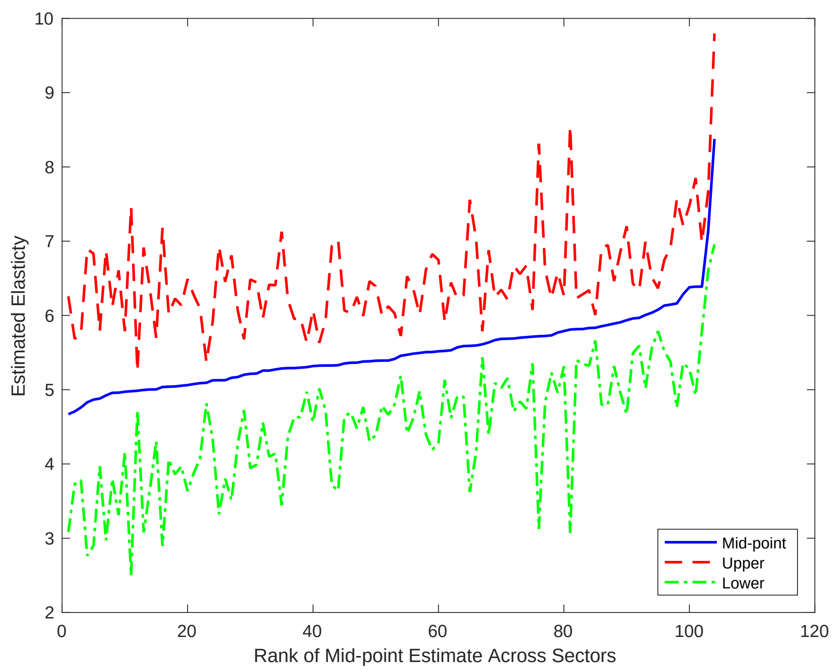
7.2 Estimates of the Elasticity of Substitution

Figure 5 shows the bounds of the distribution of our estimated elasticities of substitution for each product group at the four-quarter frequency. The median of the lower-bound elasticity is 4.62 and the median of the upper bound is 6.38, with a median midpoint of 5.40. For almost all product groups, the upper bound is the DRW estimate and the lower bound is the RW estimate, with the median RW estimate of 4.64 close to the median lower bound. This pattern of results implies a positive correlation between demand and price shocks. The upper and lower bounds identify the set of possible values of the elasticity of substitution consistent with the data for different assumptions about the correlation between demand and price shocks. The fact that these two bounds are so close together implies that, regardless of the correlation structure one assumes is present in the data, one obtains similar values for the elasticity of substitution.

[17]Even so, HomeScan data can sometimes contain entry errors. To mitigate this concern, we dropped purchases by households that reported paying more than three times or less than one third the median price for a good in a quarter or who reported buying twenty-five or more times the median quantity purchased by households buying at least one unit of the good. We also winsorized the data by dropping observations whose percentage change in price or value were in the top or bottom 1 percent.

Figure 6 displays the RW and DRW estimates (solid red and solid blue lines respectively) as well as the 95 percent point confidence intervals (dashed red and blue lines respectively).[18] Both sets of estimates are precise and we comfortably reject the null hypothesis of an implicit elasticity of 0 in Laspeyres indexes or an implicit elasticity of 1 in the Cobb-Douglas index. In other words, estimated rates of product substitution based on statistical indexes are likely to dramatically understate the degree of substitution by consumers. In terms of magnitudes, these elasticities do not differ greatly from those estimated by other studies. For example, Hottman, Redding, and Weinstein (2016), using the same data (but a different model nesting structure and estimation methodology) found that the elasticity of substitution had a median value of 3.9 across firms and 6.9 within firms. Our median estimate of 5.4, which is based on pooling the data within and across firms, falls at the midpoint between these two values.

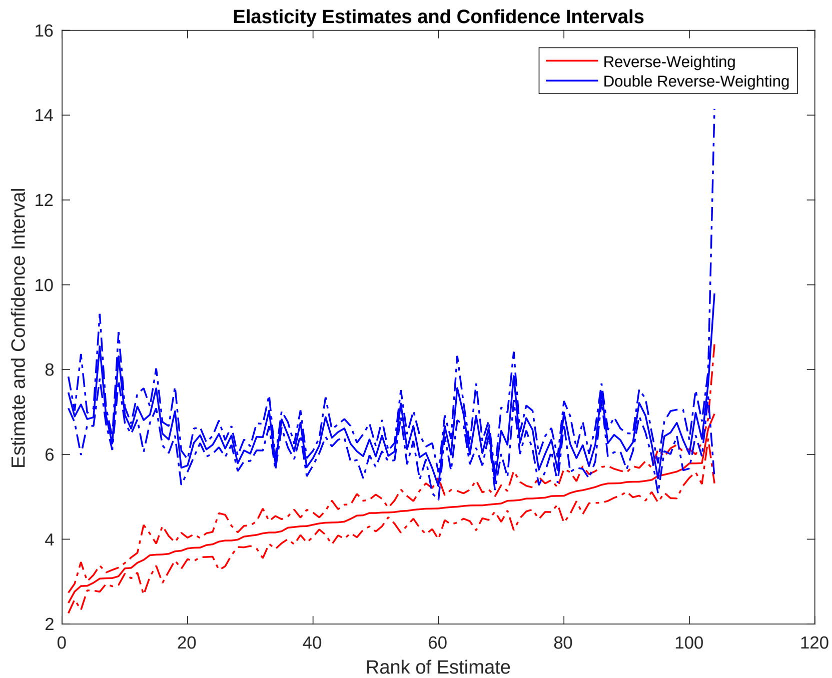
Figure 5: Distribution of RW and DRW Estimates Across Product Groups



Note: Calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

[18] We compute the confidence intervals from 50 bootstrap replications. Each bootstrap replication for a given product group resamples the observed data on the prices and expenditure shares of goods k in periods t within that product group.

Figure 6: Estimated Elasticities and Bootstrap Standard Errors



Note: Calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

To examine the importance of allowing for demand shocks for individual goods, we now impose the assumption that the demand parameters for each good are time-invariant ($\varphi_{kt} = \varphi_{kt-1} = \varphi_k$ for all $k \in \Omega_{t,t-1}$). As shown in Proposition 2, we can solve for the elasticity of substitution using the Sato-Vartia formula in this case of no demand shocks (see equation (A.12) in the web appendix). If demand shocks are small, we would expect this utility parameter to be stable as well. In order to compute how demand shocks affect the implied elasticity of substitution, we denote the implied Sato-Vartia elasticity of substitution for each period by σ_{gt}^{SV} for every four-quarter difference and product-group. We expect these estimates to vary by product group, so we are interested in the dispersion of these estimates relative to the product group mean, or $\left(\sigma_{gt}^{SV} - \frac{1}{T}\sum_t \sigma_{gt}^{SV}\right)$, where T is the number of periods. In the absence of demand shocks, we should expect this number to be zero.

Table 1 reports the mean of $\frac{1}{T}\sum_t \sigma_{gt}^{SV}$ in the first column and moments of the distribution of $\left(\sigma_{gt}^{SV} - \frac{1}{T}\sum_t \sigma_{gt}^{SV}\right)$ in the remaining columns. The mean value is 2.65 with a standard deviation of 125. Clearly, the implicit elasticities are quite volatile, and while there are some influential outliers, the volatility of the estimates permeates the distribution. Half of all observations are outside the range of 15.5 below the median implied elasticity in a product group to 12.4 above it. This enormous variation in the implied values of the elasticity of substitution, which spans all reasonable and many unreasonable values, means that the assumption of no demand shifts that underlies the Sato-Vartia formula is a problematic way of thinking about consumer behavior. If one believes the underlying assumption of the exact price index—that demand for each good is constant over time—then one must also believe that the substitution parameter between goods in the utility function varies

Table 1: Distribution of Elasticities for Each Year and Product Group

	Mean	Standard Deviation	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
Sato-Vartia Elasticity	2.65	124.87	-48.54	-15.54	-0.19	12.38	35.90
Reverse-Weighting Elasticity	4.17	1.09	-1.45	-0.62	0.06	0.74	1.24
Double Reverse-Weighting Elasticity	6.53	0.70	-0.76	-0.38	-0.03	0.34	0.71

Note: The mean elasticity is $\frac{1}{GT}\sum_{t,g}\sigma_{gt}$, and the standard deviation is the average across all product groups, g, of the standard deviation of $\left(\sigma_{gt}-\frac{1}{T}\sum_t\sigma_{gt}\right)$, where G is the number of product groups and T is the number of time periods. Percentiles correspond to the distribution of $\left(\sigma_{gt}-\frac{1}{T}\sum_t\sigma_{gt}\right)$. For the Sato-Vartia elasticity only, we exclude the top and bottom one-percent market share changes within each product group to limit the influence of outliers (including these observations results in an even higher standard deviation for the Sato-Vartia elasticity). Calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business. The mean is the average of all elasticities of substitution at the product-group level computed using equation (A.12).

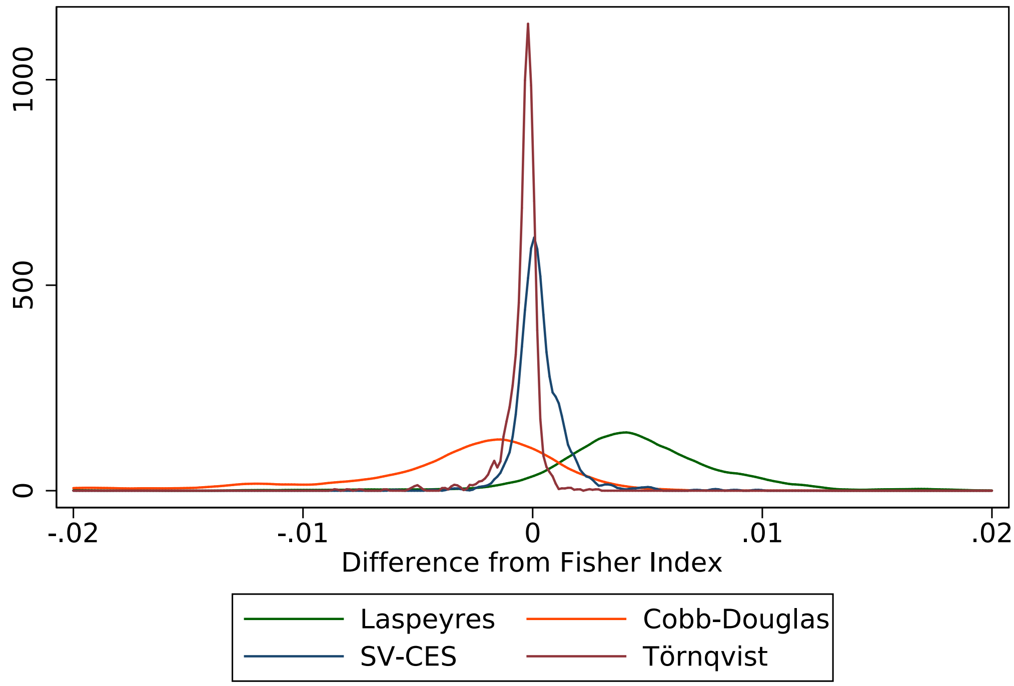
substantially over time. However, if the substitution parameter between goods varies so dramatically across pairs of periods, it is difficult to give any economic interpretation for what the price index is measuring.

Having established that assuming no demand shifts results in implausible estimates of the elasticity of substitution, we now show that our method resolves this problem. Our estimates so far pooled pairs of time periods and estimated a single elasticity of substitution (by assuming $\sigma_{gt}^{RW}=\sigma_g^{RW}$). However, theoretically, it could be the case that the elasticity of substitution is also time varying. Thus, one might wonder whether the imposition of the assumption of a common elasticity inherent in the UPI also does violence to the data. In order to see if this is is the case, we estimate σ_{gt}^{RW} for every product group and year and report the distribution of $\left(\sigma_{gt}^{RW}-\frac{1}{T}\sum_t\sigma_{gt}^{RW}\right)$ in Table 1. These estimates are much more tightly distributed around the product-group mean estimate than the time-invariant demand elasticities. The mean estimate for σ_{gt}^{RW} equals 4.2, very close to the mean value of 4.5 for σ_g^{RW}, and almost all of the annual estimates deviate from the median value for the product group by less than one. Similarly, the mean estimate of σ_{gt}^{DRW} is 6.5, the same as the mean value of 6.5 for σ_g^{DRW}. In other words, the conventional approach of assuming no demand shocks not only cannot replicate the observed expenditure shares and prices as an equilibrium of the model but also implies wildly-varying elasticities of substitution. In contrast, our unified approach exactly rationalizes the observed data on expenditure shares and prices as an equilibrium of the model for a stable elasticity of substitution. Seen in this light, the data indicates that the unified approach is the only coherent means of reconciling micro data on prices and expenditure shares with aggregate welfare measurement.

7.3 Comparison with Conventional Index Numbers

We have already argued that our framework nests many existing methods of measuring price changes and welfare. This nesting makes it possible to step-by-step show how important each assumption is in measuring

Figure 7: Each Index Differenced from the Superlative Fisher Index



Note: Calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

price changes. In general, there are three reasons why price indexes may differ: differences in the specification of substitution patterns, differences in the treatment of new goods, and differences in assumptions about whether demand shifts or not. Our next task is to quantify the importance of each of these factors. In each case, we construct price indexes for changes in the cost of living for every product group in our sample. With ten time periods and 104 product groups, we have a sample of just over 1,000 price changes.

Although the Fisher and Törnqvist indexes are not strictly nested in our setup, they are averages of the same CES building blocks.[19] In particular, the Fisher and Törnqvist indexes must lie between the bounds of the Paasche and Laspeyres indexes, two indexes that can be derived from the CES under particular parameter assumptions. The first question we need to address is how much it matters whether one uses a superlative index or a CES index. To the extent these differences are large, one might worry that adopting a CES utility function as opposed to a quadratic mean utility function (e.g., Fisher) or translog expenditure system (e.g., Törnqvist) is driving our results. While these different indexes need not be identical in theory, they are extremely similar in practice.

Figure 7 presents histograms of every four-quarter price change in our data at the product group level for each price index. We express each change in the cost of living as a difference from the superlative Fisher index, so a value of zero implies that the price index coincides with the Fisher index. The most noticeable feature

[19] All of these indexes weight price relatives using past and current expenditure shares. For example, the Törnqvist weights the log price changes by an *arithmetic* average of past and current shares while the Sato-Vartia CES index weights them by a *logarithmic* average of the two shares.

of the graph is that all of the economic indexes yield almost exactly the same changes in the cost of living in every product group in every period. The Törnqvist and Sato-Vartia CES typically record an average change in the cost of living that is identical to the Fisher index up to less than one tenth of a percentage point per year. Moreover, there is very little dispersion in these price indexes. Since the Sato-Vartia CES index is identical to the unified price index under the assumption that there are no new goods and no demand shifts for any good, we can safely say that our adoption of the CES functional form instead of a superlative index matters little for understanding changes in the cost of living. Whatever differences we find in subsequent sections must come from relaxing assumptions about the existence of demand shifts for each good or changes in the set of goods.

The fact that the CES functional form results in changes in the cost of living that are virtually identical to those of superlative indexes does not mean that *any* choice of price index yields similar results. As one can see in Figure 1, two commonly used indexes—the Cobb-Douglas and Laspeyres—are special cases of the CES in which the elasticity of substitution is one or zero, respectively. Indeed, Figure 7 shows that imposing an elasticity of zero or one on the CES functional form instead of using the Sato-Vartia formula to allow the data to dictate the implied elasticity can result in measures of cost-of-living changes that vary by around a percentage point. Nevertheless, whether one uses the CES, Fisher, or a Törnqvist matters very little in our data, so the point remains that our choice of a CES price index does not produce dramatically different results from those that would obtain from working with a translog or quadratic mean of order-r utility function.

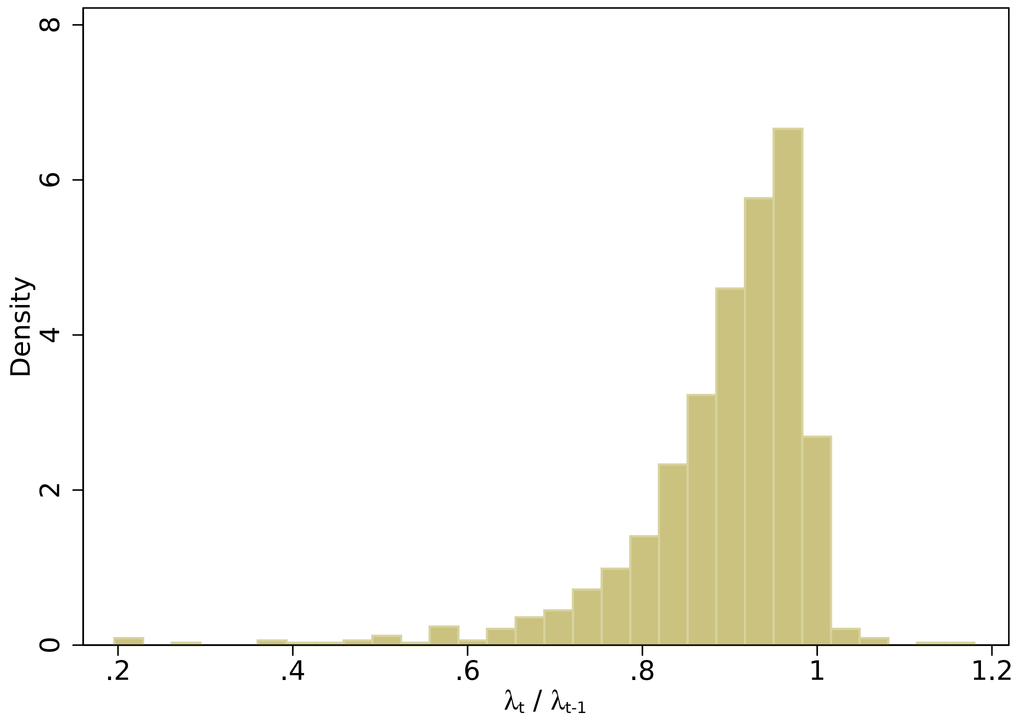
7.4 The Unified Price Index

The unified price index differs from the Sato-Vartia because it relaxes two assumptions. First, it allows for demand shifts for each good; and second, it allows for the set of goods to change over time. As we showed in Proposition 6, relaxing the first assumption gives rise to the common-goods component of the unified price index, which we know will lie below the Sato-Vartia index as long as demand shifts are positively correlated with expenditure shifts.

Relaxing the second assumption, regarding changes in the set of goods, moves us from the CG-UPI to the UPI (see equation (16)). The variety-adjustment term, which was first estimated in Feenstra (1994), combines the elasticity of substitution, which tells us how much consumers value varieties, with the rates of product creation and destruction. Figure 8 presents a histogram of the $\lambda_{t,t-1}/\lambda_{t-1,t}$ ratios that drive the variety or new-good bias. If bar codes were just turning over without upgrading, the prices and market shares of exiting bar codes would match those of new products, resulting in a $\lambda_{t,t-1}/\lambda_{t-1,t}$ ratio of one. The fact that these ratios are almost always less than one indicates that new goods tend to have lower price relative-to-demand ratios (P_{kt}/φ_{kt}) than disappearing ones. In other words, there is pervasive product upgrading.

We now examine the quantitative magnitudes of these two departures. In Figure 9, we plot the expenditure-share-weighted average of the changes in the cost of living across product groups for each of the different index numbers over time, again using the initial period expenditure share weights. It is well-known that conventional indexes—Fisher, Törnqvist and Sato-Vartia (CES)—are bounded by the Paasche and Laspeyres indexes. Thus, we can think of conventional indexes as giving us a band of cost-of-living changes that is de-

Figure 8: $\lambda_t / \lambda_{t-1}$, Four-Quarter Differences



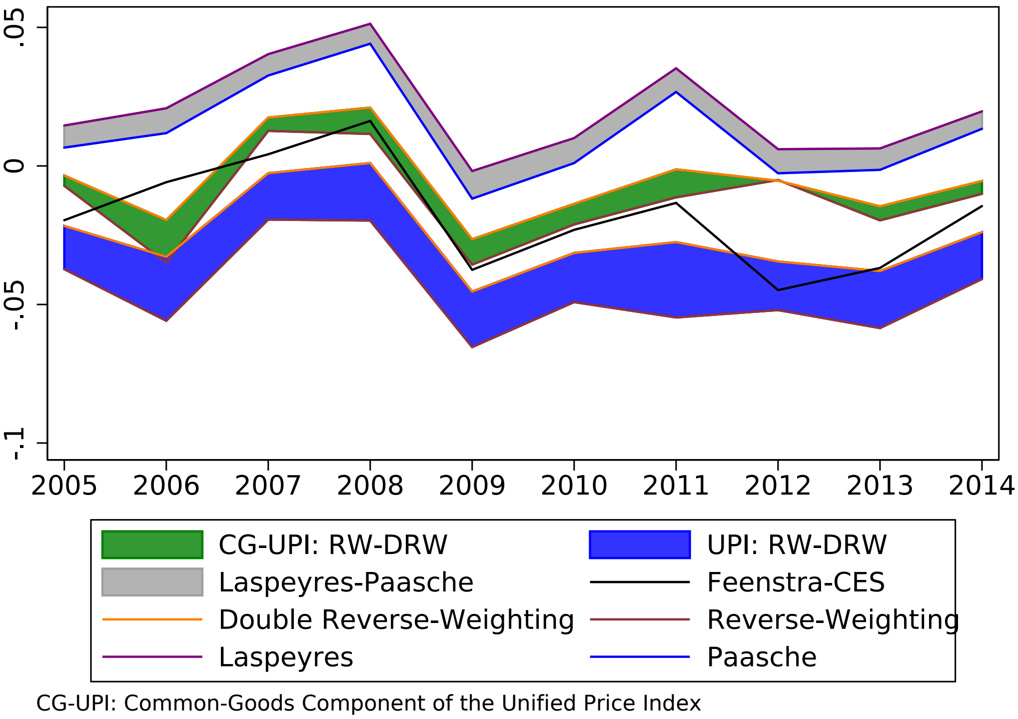
Note: Calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

termined by assumptions about consumer substitution patterns, under the assumption of no shifts in demand for any good. Similarly, the range of cost-of-living changes identified by the common-goods UPI is bounded by the different elasticities of substitution that are consistent with the data under different assumptions about the correlation between demand and price shocks. As one can see from the figure, the bands for conventional indexes and those for the CG-UPI are quite tight for each type of index, but differ across types. Therefore, assumptions about consumer substitution behavior or methods of estimating substitution elasticities are empirically much less important for understanding changes in the cost of living than assumptions about shifts in demand.

A second important difference across indexes arises from different assumptions about entry and exit. Conventional indexes ignore new and disappearing goods, which is equivalent in our setup to assuming that $\lambda_{t,t-1} / \lambda_{t-1,t} = 1$. Feenstra (1994) introduced the variety-adjustment term to the conventional (time-invariant demand) Sato-Vartia index. As we can see in the plot, adjusting a cost-of-living change for gains due to new goods is at least as important as adjusting it for demand shifts (and substantially more important than differences in the treatment of consumer substitution). The Feenstra index lies substantially below the conventional indexes, which share the no demand-shift assumption but do not allow for product turnover.

Finally, the UPI uses the Feenstra methodology for measuring new-goods bias and allows for demand shifts. The band of plausible cost-of-living changes is a bit wider than for the CG-UPI, because elasticity estimates affect both the variety adjustment and the common-goods component, but we see that the variety-

Figure 9: Aggregate Price Index, Calculated as a Share-Weighted Average Price Growth Rate



Note: Calculated based on data from The Nielsen Company (US), LLC and provided by the Marketing Data Center at The University of Chicago Booth School of Business.

adjustment substantially lowers the estimated change in the cost of living. In other words, the data indicates that when measuring changes in the cost of living, conventional indexes are biased upwards due to two counterfactual assumptions: no entry/exit and no shifts in demand for any good.

8 Conclusions

Measuring price aggregates is central to several fields of economics, including international trade and macroeconomics, but raises a number of challenges. On the one hand, micro data on prices and expenditure shares suggest that demand curves can shift with changes in preferences for individual goods. On the other hand, existing price indices that use these micro data as inputs rule out shifts in demand curves by assumption. Our first main contribution is to develop a unified approach that allows for demand shocks for individual goods (so as to rationalize the observed data) while preserving a money-metric utility function (so as to be able to make welfare comparisons over time). Our second main contribution is to develop a new method for estimating the demand parameters needed to implement this approach across many sectors for the economy as a whole. In such settings, researchers may be concerned that demand and supply shocks are correlated, but may not have access to the detailed information required to construct valid instruments for every sector. We show how assumptions about demand alone can be used to provide upper and lower bounds for the elasticity of substitution in the constant elasticity of substitution (CES) demand system.

Our central insight is to use the demand system to substitute out for the unobserved demand parameters for each good to express the change in the cost of living solely in terms of prices and expenditure shares that are directly comparable over time. We show how to implement this approach for CES and translog preferences and for a random utility model incorporating multiple types of consumers. We derive a unified price index that is exact for CES preferences and incorporates both shifts in demand for surviving goods and the entry and exit of goods over time. We show that this unified price index nests all major existing economic and statistical price indices, which can be thought of as special cases that impose parameter restrictions (e.g. an elasticity of substitution of zero), abstract from certain features of the data (the entry and exit of goods) or ignore specific implications of the model (the relationship between the demand system and the unit expenditure function). We show that abstracting from demand shocks for individual goods introduces a "consumer-valuation bias" into conventional price indices, which is analogous to the well-known substitution bias, and reflects the fact that consumers can substitute towards the goods that they like more.

To implement our unified approach for CES preferences, we develop a new "reverse-weighting" estimator for the elasticity of substitution. We show that combining the demand system with the unit expenditure function permits identification using only assumptions about demand and regardless of the correlation between demand and supply shocks. We use our reverse-weighting estimator to derive upper and lower bounds for the true elasticity of substitution. When demand and price shocks are orthogonal, these two bounds coincide, and we obtain point identification of the true elasticity of substitution. More generally, when demand and price shocks are correlated, these bounds identify a set of parameter values consistent with the observed data and the assumption of CES preferences.

We implement our unified approach using U.S. bar-code data. We show that the consumer-valuation bias in conventional price indices is substantial—as big as the new-goods bias—amounting to more than one percentage point per annum. This bias is much larger than biases arising from different elasticity estimates or assumptions about consumer substitution and suggests that conventional price indices substantially overstate the increase in the cost of living over time by assuming that demand shifts do not occur. We show that our reverse-weighting estimator identifies a narrow range of possible values for the elasticity of substitution that are consistent with the observed data and the assumption of CES preferences. Even in our most conservative specification, in which we use the upper bound for the elasticity of substitution, we find that the consumer-valuation bias remains substantial.

In conclusion, we develop a unified approach to aggregate welfare measurement that rationalizes micro data and permits exact aggregation, is easy to implement, and matters for understanding welfare.

References

[1] Anderson, James and Eric van Wincoop (2003) "Gravity with Gravitas: a Solution to the Border Puzzle," *American Economic Review*, 93(1), 170-192.

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

[3] Antràs Pol (2003) "Firms, Contracts, and Trade Structure," *Quarterly Journal of Economics*, 118(4), 1375-1418.

[4] Arkolakis, Costas and Arnaud Costinot and Andrès Rodriguez-Clare (2012) "New Trade Models, Same Old Gains," *American Economic Review*, 102(1), 94-130.

[5] Armington, Paul S. (1969) "A Theory of Demand for Products Distinguished by Place of Production," IMF Staff Papers, 16, 159-178.

[6] Atkin, David, and Dave Donaldson (2015) "Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs," *NBER Working Paper*, 21439.

[7] Atkin, David, Benjamin Faber, and Marco Gonzales-Navarro (2016) "Retail Globalization and Household Welfare: Evidence from Mexico," *Journal of Political Economy,* forthcoming.

[8] Balk, Bert M. (1999) "On Curing the CPI's Substitution and New Goods Bias," Department of Statistical Methods, Statistics Netherlands, mimeograph.

[9] Banerjee, K. S. (1983) "On the Existence of Infinitely Many Ideal Log-Change Index Numbers Associated with the CES Preference Ordering," *Statistische Hefte / Statistical Papers*, 24, 141-8.

[10] Behrens, Kristian, Yoshitsugu Kanemoto and Yasusada Murata (2014) "New Trade Models, Elusive Welfare Gains," *CEPR Discussion Paper*, 10255.

[11] Benassy, Jean-Pascal (1996) "Taste for Variety and Optimum Production Patterns in Monopolistic Competition," *Economics Letters*, 52, 41-47.

[12] Bernard, Andrew, Bradford Jensen, Jonathan Eaton and Samuel Kortum (2003) "Plants and Productivity in International Trade," *American Economic Review*, 93 (4), 1268-90.

[13] Bernard, Andrew, Stephen Redding and Peter Schott (2007) "Comparative Advantage and Heterogeneous Firms," *Review of Economic Studies*, 74(1), 31-66.

[14] Bernard, Andrew, Stephen Redding and Peter Schott (2011) "Multi-product Firms and Trade Liberalization," *Quarterly Journal of Economics*, 126(3), 1271-1318.

[15] Berry, Steven (1994) "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2), 242-262.

[16] Bils, Mark and Klenow, Peter (2001) “Quantifying Quality Growth,” *American Economic Review*, 91, 1006-1030.

[17] Blanchard, Olivier Jean and Nobuhiro Kiyotaki (1987) “Monopolistic Competition and the Effects of Aggregate Demand,” *American Economic Review*, 77(4), 647-666.

[18] Broda, Christian and David E. Weinstein (2006) “Globalization and the Gains from Variety,” *Quarterly Journal of Economics*, 121(2), 541-86.

[19] Broda, Christian and David E. Weinstein (2010) “Product Creation and Destruction: Evidence and Price Implications,” *American Economic Review*, 100, 691–723.

[20] Carli, G. (1764) “Del Valore e Della Proporzione de Metalli Monetati,” reprinted in (ed) G. G. Destefanis, *Scrittori Classici Italiani di Economia Politica*, Milan, 297-336.

[21] Caves, Douglas W., Laurits R. Christensen and W. Erwin Diewert (1982) “The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity,” *Econometrica*, 50(6), 1393-1414.

[22] Chance, W. A. (1966) “A Note on the Origins of Index Numbers,” *Review of Economics and Statistics*, 48(1), 108-110.

[23] Cobb, Charles W. and Paul H. Douglas (1928) “A Theory of Production,” *American Economic Review*, Papers and Proceedings, 18(1), 139-165.

[24] Diewert, W. E. (1976) “Exact and Superlative Index Numbers,” *Journal of Econometrics*, 4, 115-145.

[25] Dixit, Avinash and Joseph Stiglitz (1977) “Monopolistic Competition and Optimum Product Diversity,” *American Economic Review*, 297-308.

[26] Dutot, Nicolas (1738) *Réflexions Politiques sur les Finances et le Commerce*, Paris: E. Droz.

[27] Eaton, Jonathan and Kortum, Samuel (2002) “Technology, Geography, and Trade,“ *Econometrica*, 70(5), 1741-1779.

[28] Fally, Thibault and Benjamin Faber (2016) “Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data,” mimeo

[29] Feenstra, Robert C. (1994) “New Product Varieties and the Measurement of International Prices,” *American Economic Review*, 84(1), 157-177.

[30] Feenstra, Robert C. and Marshall B. Reinsdorf (2007) “Should Exact Index Numbers Have Standard Errors? Theory and Application to Asian Growth,” in (eds) Ernst R. Berndt and Charles R. Hulten, *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, Chicago: University of Chicago Press.

[31] Feenstra, Robert C. and Marshall B. Reinsdorf (2010) “An Exact Price Index for the Almost Ideal Demand System,” *Economics Letters*, 66(2) 159-162.

[32] Fisher, Irving (1922) *The Making of Index Numbers*, Boston: Houghton Mifflin.

[33] Fisher, Franklin M. and Karl Shell (1972) *The Economic Theory of Price Indices: Two Essays on the Effects of Taste, Quality, and Technological Change*, New York: Academic Press.

[34] Handbury, Jessie (2013) "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities," Wharton mimeo.

[35] Helpman Elhanan, Marc J. Melitz and Stephen R. Yeaple (2004) "Export Versus FDI with Heterogeneous Firms," *American Economic Review*, 94(1), 300-316.

[36] Hottman, Colin, Stephen Redding and David Weinstein (2016) "Quantifying the Sources of Firm Heterogeneity," *Quarterly Journal of Economics*, 131(3), 1291-1364.

[37] Hsieh, Chang-Tai and Peter Klenow (2009) "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403-1448.

[38] Jevons, William Stanley (1863) *A Serious Fall in the Value of Gold*, Edward Stanford.

[39] Jones, Charles I. and Peter J. Klenow (2016) "Beyond GDP? Welfare Across Countries and Time," *American Economic Review,* forthcoming.

[40] Klepper, Steven and Edward E. Leamer (1984) "Consistent Sets of Estimates for Regressions with Errors in All Variables," *Econometrica*, 52(1), 163-184.

[41] Konus, A. (1924) "The Problem of the True Index of the Cost of Living," translated in *Econometrica* (1939), 7, 10-29.

[42] Konyus, A. A. and S. S. Byushgens (1926) "Probleme Pokupatelnoi Cili Deneg," *Voprosi Konyunkturi*, II(1), 151-172, [English title: Conus, A.A. and S.S. Buscheguennce, 1926, On the Problem of the Purchasing Power of Money, The Problems of Economic Conditions (supplement to the Economic Bulletin of the Conjuncture Institute) 2, 151-172].

[43] Krugman, Paul R. (1980) "Scale Economies, Product Differentiation and the Pattern of Trade," *American Economic Review*, 70, 950-959.

[44] Krugman, Paul R. (1991) "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483-499.

[45] Krugman, Paul and Anthony J. Venables (1995) "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 857-880.

[46] Laspeyres, Etienne (1871) "Die Berechnung einer mittleren Warenpreissteigerung" in *Jahrbücher für Nationalökonomie und Statistik*.

[47] Lau, L. J. (1979) "On Exact Index Numbers," *Review of Economics and Statistics*, 61, 73-82.

[48] Liviatan, Nissan (1961) "Errors in Variables and Engel Curve Analysis," *Econometrica*, 29(3), 336-362.

[49] McFadden, Daniel (1974) "Conditional Logit Analysis of Qualitative Choice Behavior," in (ed) Paul Zaremba, *Frontiers in Econometrics*, New York: Academic Press.

[50] McFadden, Daniel and Kenneth Train (2000) "Mixed MNL Models of Discrete Response," *Journal of Applied Econometrics*, 15, 447-470.

[51] Melitz, Marc J. (2003) "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695-725.

[52] Moulton, B. R. (1996) "Constant Elasticity Cost-of-Living Index in Share- Relative Form," Washington DC: U. S. Bureau of Labor Statistics, mimeograph.

[53] Neary, J. Peter (2004) "Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income," *American Economic Review*, 94(5), 1411-1428.

[54] Nevo, Aviv (2003) "New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand System," *Review of Economics and Statistics*, 85(2), 266-275.

[55] Paasche, Hermann (1875) *Die Geldentwertung zu Halle in den letzten Decennien dieses Jahrhunderts*, Plötz, Halle.

[56] Sato, Kazuo (1976) "The Ideal Log-Change Index Number," *Review of Economics and Statistics*, 58(2), 223-228.

[57] Shapiro, M. D. and D. W. Wilcox, (1997) "Alternative Strategies for Aggregating Prices in the CPI," *Federal Reserve Bank of St. Louis Review*, 79(3), 113-125.

[58] Stoevska, Valentina (2008) *Consumer Price Index Manual: Theory and Practice*, Report on the 2007 Worldwide Survey on the Usefulness of the CPI Manual, Geneva: International Labor Organization (ILO).

[59] Syverson, Chad (2016) "Challenges to Mismeasurement Explanations for the U.S. Productivity Slowdown," *NBER Working Papers*, 21974.

[60] Törnqvist, Leo (1936) "The Bank of Finland Consumption Price Index," *Bank of Finland Monthly Bulletin*, 10, 1-8.

[61] Triplett, Jack E. (2001) "Should the Cost-of-living Index Provide the Conceptual Framework for a Consumer Price Index?" *Economic Journal*, 111(472), 311-334.

[62] Vartia, Yrjo (1976) "Ideal Log-Change Index Numbers," *Scandinavian Journal of Statistics*, 3(3), 121-126.

[63] Wicksell, K. (1898) *Interest and Prices*, English translation, New York 1936: Sentry Press.