

Welfare and Output with Income Effects and Taste Shocks

David R. Baqaee
UCLA

Ariel Burstein*
UCLA

October 2022

Abstract

We present a unified treatment of how welfare responds to changes in budget sets or technologies with taste shocks and non-homothetic preferences. We propose a welfare metric that ranks production possibility frontiers that differs from one that ranks budget sets, and characterize it using a general equilibrium generalization of Hicksian demand. This extends Hulten’s theorem, the basis for constructing aggregate quantity indices, to environments with non-homothetic and unstable preferences. We illustrate our results using both long- and short-run applications. In the long run, we show that if structural transformation is caused by income effects or changes in tastes, rather than substitution effects, then Baumol’s cost disease is twice as important for our preferred measure of welfare. In the short run, we show that standard chain-weighted deflators understate welfare-relevant inflation for current tastes. Finally, using the COVID-19 recession, we illustrate that chain-weighted real consumption and real GDP are unreliable metrics for measuring welfare or production when there are taste shocks.

*First version: April 2021. We thank Conor Foley, Yasutaka Koike-Mori, and Sihwan Yang for outstanding research assistance. We thank Fernando Alvarez, Andy Atkeson, David Atkin, Natalie Bau, Paco Buera, Javier Cravino, Joe Kaboski, Greg Kaplan, Stephen Redding, Esteban Rossi-Hansberg, Pierre Sarte, David Weinstein, and Jon Vogel for helpful comments. We are grateful to Emmanuel Farhi and Seamus Hogan, both of whom passed away tragically before this paper was written, for their insights and earlier conversations on these topics. This paper received support from NSF grant No. 1947611. The conclusions and analysis are our own, calculated in part on data from Nielsen Consumer LLC and provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Introduction

In this paper, we study how a change in the economic environment affects welfare. For example, how does an individual's welfare change when her budget constraint changes, or how does national welfare change when technologies change? If preferences are homothetic and there are no taste shocks, there is a well-known formula that answers these questions: the chain-weighted (Divisia) index of real consumption.¹ Assuming that there are no income effects or changes in tastes is highly convenient, but also highly inconsistent with reality. In this paper, we relax both of these assumptions and characterize changes in welfare, changes in chained-weighted consumption, and the gap between them in terms of measurable sufficient statistics.

To compare two different choice sets, indexed by t_0 and t_1 , for a consumer with preferences \succeq , we ask: "how much must the t_0 endowment change to make \succeq indifferent between the two choice sets?" If the preference relations at t_0 and t_1 are different, due to changes in tastes, then one must take a stance on which preference relation is used for the comparison because the answer depends on this choice.² In this paper, our baseline measure of welfare is the equivalent variation using t_1 preferences, though we also characterize and report results for other welfare questions (i.e. compensating variation and t_0 preferences) and explain how they differ.

We begin by studying this problem in partial equilibrium, taking prices as given, where our welfare metric compares and ranks different budget constraints. This welfare metric is appropriate for an infinitesimal agent whose choices do not alter prices. We call this a *micro* welfare metric. We then propose a generalization of money-metric measures that ranks production possibility frontiers (PPFs) and takes into account that prices are endogenous to collective choices. We call this a *macro* welfare metric. We show that comparing two PPFs using their corresponding equilibrium budget constraints, which is what standard micro money-metrics do, can be misleading if preferences are non-homothetic or unstable.

Calculating micro welfare requires integrating Hicksian (or compensated) demand curves with respect to prices. We prove that there are general equilibrium counterparts to Hicksian demand curves, which we call Hicksian *sales*, and show that integrating them with respect to technologies yields macro welfare. This general equilibrium integral also gener-

¹Chain-weighted indices weigh changes in prices or quantities by good-specific weights that are updated every period. The continuous time analog is called a *Divisia* index. Chained-weighted indices are the standard way to measure real economic activity (real GDP, productivity, real consumption, etc.), and their use is justified under homotheticity and stability of preferences (see e.g. Chapters 15 in the UN's System of National Accounts Manual, United Nations 2009, and Chapters 15 and 17 in the CPI Manual, IMF 2004).

²For background on how to account for taste shocks in welfare measures, see Fisher and Shell (1968) and Samuelson and Swamy (1974).

alizes Hulten’s theorem to measure welfare in environments featuring taste shocks and non-homotheticities. In this paper, we focus on neoclassical economies with homogeneous agents, but a companion paper (Baqee and Burstein, 2021) generalizes our results to economies with heterogeneous agents and distortions.

We provide exact and approximate characterizations of the change in micro and macro welfare. In contrast to the standard chain-weighted consumption index formula, which weighs changes in prices or technologies using observed shares, welfare-relevant indices weigh changes in prices or technologies using Hicksian shares. We show that compared to our baseline welfare measure, this implies that that chain-weighted consumption *undercounts* expenditure switching due to income effects or taste shocks (but not substitution effects).

To understand why chain-weighted indices undercount expenditure-switching due to income effects and taste shocks, consider the following example. Over the post-war period, spending on healthcare grew relative to manufacturing. Since older and richer consumers spend more on healthcare, suppose this was caused by consumers getting older and richer. In this case, a chain-weighted consumption index does not correctly account for expenditure-switching by consumers. Intuitively, when we compare the past to the present, we must use demand curves that are relevant for the older and richer consumers of today, and not demand curves that were relevant in the past. Whereas a chained deflator weighs changes in prices that happened during the 1950’s using demand from the 1950’s, a welfare-relevant index uses demand from today to weigh changes in prices throughout the sample. We show that the chain-weighted consumption index is higher than the welfare-relevant index if income- or taste-driven expenditure-switching is positively correlated with changes in prices.

Our results for welfare and the gap between welfare and real consumption are expressed in terms of measurable sufficient statistics. In both partial and general equilibrium, we show that computing the change in welfare does not require direct knowledge of the taste shocks or income elasticities. Instead, what we must know are expenditure shares and elasticities of substitution at the final allocation. For micro welfare, these are the household’s expenditure shares and elasticities of substitution in consumption. For macro welfare, these are the input-output table and elasticities of substitution in production and consumption. Our results can be used both for ex-post accounting and ex-ante counterfactuals.

For very simple economies with one factor, constant returns to scale, and no intermediates, the difference between welfare and chain-weighted consumption is approximately half the covariance of supply and demand shocks. We generalize this formula to more com-

plex economies and show how the details of the production structure, like input-output linkages, complementarities in production, and decreasing returns to scale, can interact with non-homotheticities and preference shocks to magnify the gap between welfare and chain-weighted consumption. Even if there are only idiosyncratic microeconomics shocks, the discrepancies between welfare and chain-weighted real consumption that we emphasize do not average out at the aggregate level.

We illustrate the relevance of our results for understanding long-run and short-run phenomena with three applications.

- i. **Long-run application:** Baumol (1967) argues that aggregate productivity growth slows down if industries with relatively low productivity growth become larger as a share of the economy over time. To be specific, from 1947 to 2014, aggregate total factor productivity (TFP) in the United States grew by 60%. If the U.S. economy had kept its original 1947 industrial structure, then aggregate TFP would have grown by 78% instead. We show that if this transformation is caused solely by income effects and demand instability, rather than substitution effects, then our baseline measure of welfare-relevant TFP grew by only 47% instead of 60%.
- ii. **Short-run application:** Our second application looks at shorter horizons but uses more disaggregated (and volatile) data. We compute changes in the welfare-relevant price index and compare these to the chained index using product-level non-durable consumer goods data between 2004 and 2019. We find that the chained-weighted price index understates inflation rates if we use 2019 preferences. This is because goods that became more popular over time, due to taste shocks, have higher inflation rates. At annual frequency, the gap is around 1 percentage point, and it grows to around 4 percentage points over the whole sample. This also means that for past preferences, rather than current preferences, the welfare-relevant inflation rate is lower than the chained index because goods that became less popular experienced lower inflation.
- iii. **Business-cycle application:** Our final application draws on the COVID-19 recession to illustrate the difference between macro and micro notions of welfare. During this recession, household expenditures switched to favor certain sectors at the same time that those sectors experienced higher inflation. We show this implies that micro welfare using mid-2020 preferences, taking changes in prices as given, fell by more than macro welfare, taking into account that changes in prices are themselves caused by demand shocks. Furthermore, unlike macro welfare, real consumption, real GDP, and aggregate TFP are unreliable metrics for measuring changes in productive capacity

because they depend on irrelevant details like the order in which supply and demand shocks hit the economy.

In addition to preference stability and homotheticity, a chained index accurately measures welfare only if prices and quantities change continuously and are measured correctly. Many of the well-known reasons chained indices fail to measure welfare are due to violations of these measurement assumptions. For example, it is well known that real consumption fails to account for product creation and destruction if we do not measure the quantity of goods continuously as their price falls from or goes to their choke price (Hicks, 1940; Feenstra, 1994; Hausman, 1996; Aghion et al., 2019); real consumption does not properly account for changes in the quality of goods (see Syverson, 2017); and, real consumption fails to properly account for changes in non-market components of welfare, like changes in the user-cost of durable consumption or leisure and mortality (see Jones and Klenow, 2016). In these cases, the problem is that some of the relevant prices or quantities in the consumption bundle are missing or mismeasured, and correcting the index involves imputing a value for these missing prices or quantities. Non-homotheticities and taste shocks are different from mismeasured prices because they violate the maintained assumptions about preferences, not prices, and correcting the index requires using welfare-relevant (rather than observed) expenditure shares. For this reason, we abstract from these mismeasurement issues and assume that prices and quantities have been correctly measured. If prices and quantities are mismeasured or missing, then our results would apply to the quality-adjusted, corrected version of prices instead of observed prices.

Other related literature. Measuring changes in welfare using money-metrics is standard in microeconomic theory (see, e.g. chapter 7 in Deaton and Muellbauer, 1980). We characterize the gap between this notion of welfare and real consumption with non-homotheticities and taste shocks. Our general equilibrium results relax the standard assumption in growth accounting that there exists a stable and homothetic final good aggregator (extending Domar 1961; Hulten 1978). This is also an important maintained assumption in the literature on disaggregated and production network models of the economy (see Carvalho and Tahbaz-Salehi, 2019 and the references therein).

Our approach focusing on the money-metric at final preferences can be contrasted with common practice in the literature on index numbers, which focuses on Konüs price indices for intermediate levels of utility or tastes between t_0 and t_1 (see Diewert 1976, Caves et al. 1982, and Feenstra and Reinsdorf 2007). These papers show that, under some assumptions (i.e. translog or constant elasticity of substitution, CES), commonly used indices like Tornqvist and Sato-Vartia do answer a meaningful question. The advantage of this approach is

that constructing these indices requires far less information; the disadvantage is that, unlike our preferred welfare measure, these indices are not money-metrics that can be used for policy or counterfactual analysis, and they do not provide specific information about the reference indifference curve being used (what budget level, price vector, and preferences it corresponds to). Furthermore, in practice, most index numbers are constructed by chaining, and the aforementioned results do not apply to chained indices. An additional contribution of our paper is to characterize how welfare measures differ from chained (Divisia) indices.³ Finally, relative to this literature, we also provide a unified analysis of non-homotheticity and taste shocks, and we define and characterize a welfare measure for comparing PPFs rather than budget sets (taking into account that prices are endogenous to choices).

A recent and related paper is Redding and Weinstein (2020), who also study welfare changes with taste shocks and non-homotheticities. Their approach depends on untestable assumptions about cardinal properties of utility functions. On the other hand, we use a money-metric approach that relies only on ordinal preference relations and not on the way utility is cardinalized. We compare our approach to that of Redding and Weinstein (2020) in Section 2.4.

Our approach to calculate ex-post welfare changes requires knowing price changes. When information on prices is incomplete, if preferences are non-homothetic an alternative approach is to infer ex-post changes in welfare by relying on changes in prices, expenditures, price elasticities, and Engel curve slopes for only a subset of goods, given assumptions on separability and stability in preferences (see Hamilton, 2001 and more recently Atkin et al., 2020). In a different vein, Jaravel and Lashkari (2022) provide a procedure to measure microeconomic welfare changes without direct knowledge of elasticities of substitution in the absence of taste shocks and under some additional assumptions.⁴ We are interested in ex-post welfare measurement with non-homotheticities, but unlike the aforementioned papers, we are also interested in taste shocks, counterfactuals, and general equilibrium.

Our paper is also related to the literature on structural transformation and Baumol's

³Under the assumption that the path of prices is linear, Feenstra and Reinsdorf (2000) show the equivalence between Divisia and a Konüs price index for an intermediate utility level under AIDS preferences. In practice, price paths tend to be nonlinear (for evidence using scanner-level data, see Ivancic et al., 2011). Therefore, in contrast to Tornqvist and Sato-Vartia, chained indices cannot generically be interpreted as welfare measures corresponding to any well-defined preferences. This is because, as we discuss in Section 5, Divisia (or chained) indices are path-dependent, so they can violate basic properties like assigning a higher value to a strictly larger choice set. Oulton (2008) discusses how Konüs price indices resolve the path-dependency problem of Divisia indices.

⁴Baqae et al. (2022) show how Lemma 1 in this paper can also be used to accomplish this task under similar assumptions.

cost disease. As explained by Buera and Kaboski (2009) and Herrendorf et al. (2013), this literature advances two microfoundations for structural transformation. The first explanation is all about relative prices differences: if demand curves are not unit-price-elastic, then changes in relative prices change expenditure shares (e.g. Ngai and Pissarides, 2007; Acemoglu and Guerrieri, 2008; Buera et al., 2015). The second explanation emphasizes shifts in demand curves caused by income effects or taste shocks— households spend more of their income on some goods as they become richer (e.g. Kongsamut et al., 2001; Boppart, 2014; Comin et al., 2021; Alder et al., 2022) or older (Cravino et al., 2019). Our results suggest that structural transformation driven by relative price changes has different welfare implications than structural transformation driven by non-homotheticity or taste changes.

The structure of the paper is as follows. In Section 2, we set up the microeconomic problem and provide exact and approximate characterizations of welfare and chain-weighted real consumption. In Section 3, we set up the macroeconomic general equilibrium model and provide exact and approximate characterizations of welfare and chain-weighted real output changes. Whereas in Section 3 we present our macro results in terms of endogenous sufficient statistics, in Section 4 we solve for these endogenous sufficient statistics in terms of microeconomic primitives and consider some simple but instructive analytical examples. Our applications are in Section 5. We discuss how to use our results in dynamic settings and how to treat new goods in Section 6, and we conclude in Section 7. Proofs are in the appendix and supplementary materials are in an online appendix.

2 Microeconomic Changes in Welfare and Consumption

In this section, we consider changes in budget constraints in partial equilibrium. We ask how consumers value these changes, and compare this with chain-weighted real consumption. This section builds intuition for Section 3, where we model the equilibrium determination of prices.

2.1 Environment and Definitions

Consider a set of preference relations, $\{\succeq_x: x \in X\}$, over bundles of goods $c \in \mathbb{R}^N$ where N is the number of goods. The vector c includes all relevant goods, and if \succeq_x is intertemporal, then c is a path of current and future consumption bundles. These preferences are indexed by some vector of parameters, denoted by x , that the consumer does not make choices about but that can affect preference rankings over bundles of goods. For example, x could

include calendar time, age, exposure to fads, or state of nature.⁵

For every $x \in X$, we represent the preference relation \succeq_x by a utility function $U(c; x)$. Because the consumer makes no choices over x , preferences over x , if they exist, are not revealed by choices. Hence, whereas $U(c; x) > U(c'; x)$ if, and only if, $c \succeq_x c'$, a comparison of $U(c; x')$ and $U(c; x)$ is not meaningful and does not encode any information, because it is affected by how $U(\cdot; x)$ and $U(\cdot; x')$ are cardinalized.

Note that x should not be interpreted as quality because consumers have preferences and choices over quality but not over parameters of their own utility function (tastes). Hence, the welfare implications of quality changes are different to those of taste changes. Moreover, quality has interpretable cardinal units (e.g., GHz for the processor of a computer) whereas tastes do not. Quality characteristics must be included as a part of the description of the consumption bundle c , not in x . See Online Appendix C for a more detailed discussion.

There are two properties of preferences that are analytically convenient benchmarks throughout the rest of the analysis.

Definition 1. Preferences are *homothetic* if whenever $c \sim_x c'$ then $\alpha c \sim_x \alpha c'$ for every $\alpha > 0$.

When \succeq_x is homothetic, we can write $U(c; x)$ so that for every $\alpha > 0$, $U(\alpha c; x) = \alpha U(c; x)$.

Definition 2. Preferences are *stable* if \succeq_x is the same as $\succeq_{x'}$ for every x and x' in X .

If preferences are stable, then the utility function $U(c; x)$ is separable in c and x .

The indirect utility function, for any value of x , is

$$v(p, I; x) = \max_c \{U(c; x) : p \cdot c = I\},$$

where p is a price vector over goods and I is expenditures (which we interchangeably refer to as income).

Consider shifts in the budget set as prices and income change from p_{t_0} and I_{t_0} to p_{t_1} and I_{t_1} . Here, t_0 and t_1 simply index the vector of prices and income being compared. Motivated by our applications, we refer to this index as time. This change in the budget set is accompanied by changes in preferences from x_{t_0} to x_{t_1} .

Because changes in utility do not have meaningful units, we use a money-metric to measure how the consumer values different budget sets. Our baseline measure of microeconomic welfare is defined as follows.

⁵For static comparisons aging can be represented using x . This is not the case in a dynamic life-cycle model, where c is the entire path of consumption over the life-cycle. In this case, preferences over consumption paths are stable even if age affects intratemporal choices (e.g. healthcare versus food when you are young or old). Therefore, standard welfare results with stable preferences apply for ranking price paths over the life-cycle.

Definition 3 (Micro Welfare). The change in welfare measured using the *micro equivalent variation* with *final preferences* is $EV^m(p_{t_0}, I_{t_0}, p_{t_1}, I_{t_1}; x_{t_1}) = \phi$ where ϕ solves

$$v(p_{t_1}, I_{t_1}; x_{t_1}) = v(p_{t_0}, e^\phi I_{t_0}; x_{t_1}). \quad (1)$$

In words, EV^m is the change in income (in logs), under initial prices p_{t_0} , that a consumer with preferences $\succeq_{x_{t_1}}$ would need to be indifferent between the budget set defined by initial prices $(p_{t_0}, e^\phi I_{t_0})$ and the new budget set defined by new prices and income (p_{t_1}, I_{t_1}) . The new budget set (p_{t_1}, I_{t_1}) is preferred to the initial one (p_{t_0}, I_{t_0}) if, and only if, EV^m is positive.

As Fisher and Shell (1968) point out, this is different to the following question that one may wish to answer: “how much better off is the consumer in t_1 compared to the consumer in t_0 ?” This question is unanswerable using only ordinal choice information. As discussed above, a comparison of two different utility functions, $U(\cdot; x_{t_0})$ and $U(\cdot; x_{t_1})$, depends on how each utility function cardinalizes the underlying preference relation, which is arbitrary because utility functions are defined only up to monotone transformations. Thus, a welfare measure based only on the ordinal preference relation must hold preference parameters x constant in the comparison. For a discussion of papers that adopt a cardinal approach, see Section 2.4.

Other than keeping x constant, we make two other choices in Definition 3. First, we focus on *final* rather than *initial* preferences, and second, we use *equivalent* rather than *compensating* variation. In principle, one could study initial preferences and compensating variation instead. In general, since these alternative measures ask different questions they give different answers, unless preferences happen to be both stable and homothetic. Our methods can be also used to understand these alternative welfare measures, but to streamline the exposition we leave some of the additional details for these alternative measures in Online Appendix A.

In our baseline, we focus on final preferences $\succeq_{x_{t_1}}$, as opposed to initial preferences $\succeq_{x_{t_0}}$, because for temporal comparisons, the asymmetry of time makes current preferences more relevant than preferences in the past. As Fisher and Shell (1968, page 5) write, “...every practical question which one wants the cost of living index to answer is answered with reference to current, not base-year tastes.” Of course, in some contexts one may be interested in using past preferences to value choice sets, and for counterfactuals one may want to use current rather than future preferences.⁶

⁶For optimal policy questions in societies with heterogeneous and/or changing preferences, one must specify a social welfare function that makes the necessary interpersonal utility comparisons. However, our paper is about assigning value to choices given some preferences, not about how different preferences should

We also focus on equivalent variation as our benchmark, rather than compensating variation. We focus on equivalent variation because, unlike the compensating variation, it is a money-metric (see McKenzie and Pearce, 1982). Specifically, the equivalent variation is itself an index of utility which transforms the utility value of different outcomes into dollar values under a common price system (prices in t_0). This is not true for compensating variation, as discussed in Section 2.4.

We define real consumption based on a chain-weighted index. This is the standard procedure used by national income accountants and statistical agencies to construct aggregate quantities given data on the evolution of prices p and consumption bundles c .

Definition 4 (Real consumption). For some path of prices and quantities that unfold as a function of time t , the change in *real consumption* from t_0 to t_1 is defined to be

$$\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} b_{it} \frac{d \log c_{it}}{dt} dt = \int_{t_0}^{t_1} \sum_{i \in N} b_i d \log c_i, \quad (2)$$

where $b_{it} \equiv p_{it}c_{it}/I_t$ is the budget share of good i given prices, income, and preferences at time t .⁷

Equation (2) is called a *Divisia* quantity index. In practice, because perfect data are not available in continuous time, statistical agencies approximate this integral via a (Riemann) sum, which they call a *chained* index. We abstract from the imperfections of these discrete time approximations in this paper.⁸ Moreover, we assume that the data on prices and quantities is *perfect* — completely accurate, comprehensive, and adjusted for any necessary quality changes. This is because the important and well-studied biases between measured and welfare-relevant objects associated with imperfections in the data, like the lack of quality adjustment, missing prices, or infrequent measurement, are different to the biases we study.

be weighed against each other.

⁷For any variable z , we use the shorthand dz to denote its change over infinitesimal time intervals, $dz \equiv \frac{dz}{dt} dt$, so that $\Delta z = \int_{t_0}^{t_1} dz$. The last term on the right-hand side of (2) suppresses dependence on t in the integral. We sometimes use this convention to simplify notation.

⁸In discrete time, one can approximate this Riemann integral in different ways. For example, we can use left-Riemann sums (Chained Laspeyres), right-Riemann sums (Chained Paasche), or mid-point Riemann sums (Chained Tornqvist or Fisher). In continuous time, all of these procedures are equivalent and yield the same answer.

2.2 Characterization of Welfare

To characterize changes in welfare, define the expenditure function for any value of x to be

$$e(p, u; x) = \min_c \left\{ \sum_{i \in N} p_i c_i : U(c; x) = u \right\}.$$

The compensated or Hicksian budget share of good i (given prices p , preferences x , and a level of utility u) is

$$b_i(p, u; x) \equiv \frac{p_i c_i(p, u; x)}{e(p, u; x)} = \frac{\partial \log e(p, u; x)}{\partial \log p_i}, \quad (3)$$

where $c_i(p, u; x)$ is Hicksian demand for good i and preferences x , and the second equality is Shephard's lemma.

The following lemma expresses changes in welfare given changes in prices and income. It is a key result that we use throughout the paper.

Lemma 1 (Micro Welfare). *For any smooth path of prices, income, and tastes that unfold as a function of time t , micro welfare changes are given by*

$$EV^m = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} b_i(p_t, u_{t_1}, x_{t_1}) \frac{d \log p_{it}}{dt} dt, \quad (4)$$

where $b_i(p_t, u_{t_1}, x_{t_1})$ denotes Hicksian budget shares at prices p_t fixing final preferences x_{t_1} and final utility $u_{t_1} = v(p_{t_1}, I_{t_1}; x_{t_1})$.

Lemma 1 follows from the observation that EV^m can be re-expressed, using the expenditure function, as

$$EV^m = \log \frac{e(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}{e(p_{t_0}, v(p_{t_0}, I_{t_0}; x_{t_1}); x_{t_1})} = \Delta \log I - \log \frac{e(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}{e(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}, \quad (5)$$

and recognizing that the second term can be written as the integral in (4).^{9,10}

The budget shares $b_i(p_t, u_{t_1}, x_{t_1})$ that are used to calculate welfare in equation (4) can be reinterpreted as those under prices p_t of a fictional consumer with homothetic and stable preferences with expenditure function $\tilde{e}(p, u) = e(p, u_{t_1}; x_{t_1}) u$. This implies that to com-

⁹The ratio of expenditure functions, holding fixed utility, is called a Konüs (1939) price index. Equation (5) shows that EV^m requires deflating nominal income by the t_1 utility and preferences Konüs price index. This is in contrast to common practice in the index number theory literature, say Diewert (1976), that uses an intermediate level of utility or preferences between u_{t_0} and u_{t_1} . We discuss how our results relate to this alternative approach in Online Appendix B.

¹⁰By definition, EV^m depends only on initial and final prices and income, given t_1 preferences. Therefore, the integral in (4) must be path-independent and yields the same answer under every continuously differentiable path of prices that go from p_{t_0} to p_{t_1} .

pute $b_i(p_t, u_{t_1}, x_{t_1})$ and EV^m , we need to know the budget shares and elasticities of substitution at t_1 but not the income elasticities or the demand shocks. This holds generally, but we illustrate it using the convenient non-homothetic CES functional form below.

Example 1 (Non-homothetic CES.). To illustrate how Lemma 1 can be used, consider a non-homothetic CES example as in Hanoch (1975), Comin et al. (2021), Matsuyama (2019), and Fally (2022). The expenditure function can be written as

$$e(p, u; x) = \left(\sum_{i \in N} x_i p_i^{1-\theta_0} u^{\xi_i} \right)^{\frac{1}{1-\theta_0}}, \quad (6)$$

with budget shares

$$b_i(p, u, x) = \frac{x_i u^{\xi_i} p_i^{1-\theta_0}}{\sum_{j \in N} x_j u^{\xi_j} p_j^{1-\theta_0}}. \quad (7)$$

The parameter θ_0 is the elasticity of substitution across goods and ξ_i governs the utility elasticity of good i . When the utility elasticities ξ_i are equal for every i , preferences are homothetic. When the set of preference parameters $\{x_i\}$ only scale proportionately, preferences are stable. Note that the underlying preference relation only pins down relative ξ_i 's and relative x_i 's, but not their overall levels. That is, scaling all x_i by the same positive number and adding the same constant to every ξ_i has no effect on choices, but alters the utility associated with each choice.

Welfare-relevant budget shares $b(p, u_{t_1}, x_{t_1})$ as functions of prices p can be calculated as

$$b_i(p, u_{t_1}; x_{t_1}) = \frac{b_{it_1} \times (p_i / p_{it_1})^{1-\theta_0}}{\sum_{j \in N} b_{jt_1} \times (p_j / p_{jt_1})^{1-\theta_0}}. \quad (8)$$

Using (8) in the integral in Lemma 1 yields

$$EV^m = \Delta \log I + \log \left(\sum_{i \in N} b_{it_1} \left(\frac{p_{it_0}}{p_{it_1}} \right)^{1-\theta_0} \right)^{\frac{1}{1-\theta_0}}. \quad (9)$$

Hence, to compute $b(p, u_{t_1}, x_{t_1})$ and EV^m , we need to know the final budget shares b_{t_1} and the elasticity of substitution θ_0 , but not the utility elasticities ξ or taste shocks x .¹¹ Equation (9) shows when CES preferences are non-homothetic or unstable, the price deflator for EV^m

¹¹In practice, estimating the elasticity of substitution θ_0 must take into account the possibility of demand shocks and income effects. For example, Auer et al. (2022) estimate compensated price elasticities and apply Lemma 1 to measure the heterogeneous welfare effects of changes in foreign prices in the presence of demand non-homotheticities.

is the exact hat-algebra CES deflator (sometimes referred to as the Lloyd-Moulton index) from t_1 to t_0 rather than the Sato-Vartia index.

If b_{t_0} is known but b_{t_1} is not, we first have to calculate b_{t_1} before applying (9). Online Appendix D shows how this can be done given knowledge of taste shocks and income elasticities of demand.

2.3 Comparing Welfare and Real Consumption

Using the budget constraint, real consumption in (2) can be expressed in terms of changes in nominal income deflated by price changes:

$$\Delta \log Y = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} b_i(p_t, u_t, x_t) \frac{d \log p_{it}}{dt} dt, \quad (10)$$

where u_t is $v(p_t, I_t; x_t)$ and $b(p_t, u_t, x_t)$ are observed budget shares, b_{it} , at t . In words, changes in real consumption are equal to changes in income minus changes in the consumption price deflator. Changes in real consumption (and the consumption price deflator) potentially depend on the entire path of prices and quantities between t_0 and t_1 and not just the initial and final values. This is unlike welfare changes, EV^m , which depend only on initial and final prices and incomes and not on their entire path.

Combining (4) and (10) implies the following proposition.

Proposition 1 (Micro Welfare and Real Consumption). *Given a smooth path of prices, income, and tastes that unfold as a function of time t , the difference between welfare changes and real consumption is*

$$EV^m - \Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} (b_i(p_t, u_t, x_t) - b_i(p_t, u_{t_1}, x_{t_1})) \frac{d \log p_{it}}{dt} dt. \quad (11)$$

This proposition clarifies the differences between welfare and real consumption. Real consumption weighs changes in prices at time t by observed budget shares at time t , taking into account expenditure-switching as it happens. In contrast, welfare takes into account expenditure-switching due to income effects and taste shocks from the beginning, weighing changes in prices at time t by $b_i(p_t, u_{t_1}, x_{t_1})$. Intuitively, EV^m depends on budget shares evaluated at final utility (u_{t_1}) and tastes (x_{t_1}), since EV^m adjusts the level of income in t_0 to make consumers with t_1 preferences as well off as they are in t_1 . For instance, if the consumer prefers the t_1 to the t_0 budget set, then she must be given more income in t_0 to make her indifferent between t_0 and t_1 . As we give her more income in t_0 , the shape of her indifference curve changes until it mirrors the one in t_1 . This means that the shape of the

indifference curve relevant for the comparison is the one at t_1 .¹²

An immediate consequence of Proposition 1 is the well-known result that real consumption is equal to changes in equivalent variation if, and only if, preferences are homothetic and stable. This is because when preferences are stable and homothetic, budget shares do not depend on x or on utility u over time. That is, when preferences are homothetic and stable, $b_{it} = b_i(p_t, u_t, x_t) = b_i(p_t, u_{t_1}, x_{t_1})$ for every path of shocks and every t . Real consumption and welfare are also the same if the difference between the actual and Hicksian budget shares in equation (11) are orthogonal to price changes.

To gain more intuition, we provide a second-order approximation as the time period goes to zero, $t_1 - t_0 = \Delta t \rightarrow 0$. Formally, for any variable z , we write $\Delta \log z \approx (\partial \log z / \partial t) \Delta t + 1/2(\partial^2 \log z / \partial t^2)(\Delta t)^2$, where the remainder in the approximation is of order $(\Delta t)^3$.¹³ Furthermore, throughout we use the shorthand $\mathbb{E}_z(y)$ for $\sum_i z_i y_i$ for any pair of vectors z and y where $z_i \geq 0$ and $\sum_i z_i = 1$. Similarly, $Cov_z(w, y)$ is shorthand for $\mathbb{E}_z(wy) - \mathbb{E}_z(w)\mathbb{E}_z(y)$.

Proposition 2 (Approximate Real Consumption and Micro Welfare). *To a second-order approximation, the change in real consumption is*

$$\Delta \log Y \approx \Delta \log I - \mathbb{E}_{b_{t_0}}(\Delta \log p) - \underbrace{\frac{1}{2} Cov_{b_{t_0}}(\Delta \log b, \Delta \log p)}_{\text{expenditure-switching}}, \quad (12)$$

and the change in welfare is

$$EV^m \approx \Delta \log I - \mathbb{E}_{b_{t_0}}(\Delta \log p) - Cov_{b_{t_0}}(\Delta \log b, \Delta \log p) + \underbrace{\frac{1}{2} \sum_{i,j} \Delta \log p_j \frac{\partial b_i}{\partial \log p_j} \Delta \log p_i}_{\substack{\text{expenditure-switching} \\ \text{due to substitution effect}}}, \quad (13)$$

where $\partial b / \partial \log p$ is the partial derivative of the Hicksian budget share with respect to prices p evaluated at t_0 .

We discuss real consumption and welfare in turn. The first two terms in (12) are just the change in nominal income deflated by average price changes. The last term (multiplied

¹²When there are no taste shocks, real consumption, defined by (2), is a multi-good version of the change in consumer surplus, which is the area under the Marshallian demand curve. Similarly, by equation (4), welfare is the area under a Hicksian demand curve. Hence, in a partial equilibrium context with stable preferences, the gap between real consumption and welfare is also the gap between consumer surplus and welfare, studied by Hausman (1981) and McKenzie and Pearce (1982) amongst others. This equivalence does not hold when preferences are unstable because in this case Marshallian consumer surplus is not the same as chained real consumption.

¹³For local approximations throughout the paper, we assume that the exogenous parameters are smooth functions of t and that the expenditure function is a smooth function of preference parameters x .

by 1/2) captures how expenditure-switching affects real consumption. If expenditures rise for goods whose prices are rising, then this lowers real consumption. In principle, the expenditure shares can change for three different reasons: substitution effects, income effects, and taste shocks.

Equation (13) shows that the first-order terms in welfare are identical to real consumption, but discrepancies are present at the second order. Specifically, welfare treats changes in budget shares due to substitution effects differently to changes in budget shares due to taste shocks or income effects. The adjustment term in (13) implies that welfare puts more weight on expenditure-switching if it is due to income effects or taste shocks than if it is due to substitution effects. To better understand the expenditure-switching terms, consider the non-homothetic CES demand system. To apply Proposition 2, we need to approximate changes in expenditure shares only to the first-order. This is because $\Delta \log b$ is always multiplied by $\Delta \log p$, so any nonlinear terms in $\Delta \log b$ will have order higher than two once they are multiplied by $\Delta \log p$. Log-linearizing the budget shares (7) with respect to time yields¹⁴

$$d \log b_i = \underbrace{[1 - \theta_0] [d \log p_i - \mathbb{E}_b(d \log p)]}_{\text{substitution effects}} + \underbrace{[\varepsilon_i - 1] [d \log I - \mathbb{E}_b(d \log p)]}_{\text{income effects}} + \underbrace{d \log x_i}_{\text{taste shifter}}, \quad (14)$$

where ε is the vector of income elasticities of demand for each good, and $d \log x$ is the vector of demand shifters due to taste shocks.¹⁵ Substituting (14) into Proposition 2 results in the following proposition.

Proposition 3 (Approximation with Non-Homothetic CES). *For a consumer with non-homothetic preferences, defined by (6), to a second-order approximation, the change in real consumption is*

$$\begin{aligned} \Delta \log Y \approx & \Delta \log I - \mathbb{E}_{b_{t_0}}(\Delta \log p) - \underbrace{\frac{1}{2}(1 - \theta_0) \text{Var}_{b_{t_0}}(\Delta \log p)}_{\text{expenditure-switching due to substitution effect}} \\ & - \frac{1}{2} \underbrace{\text{Cov}_{b_{t_0}}(\Delta \log x, \Delta \log p)}_{\text{expenditure-switching due to taste shock}} - \frac{1}{2} \underbrace{\left[\Delta \log I - \mathbb{E}_{b_{t_0}}(\Delta \log p) \right] \text{Cov}_{b_{t_0}}(\varepsilon, \Delta \log p)}_{\text{expenditure-switching due to income effect}}, \end{aligned} \quad (15)$$

¹⁴See Online Appendix D for a more detailed derivation. We assume here that all goods have positive expenditure shares. We discuss new goods in Section 6.

¹⁵In terms of primitives of the utility function, $\varepsilon_i = 1 + (1 - \theta_0)(\xi_i / \mathbb{E}_b(\xi) - 1)$ and $d \log x_i = d \log x_i - (\xi_i / \mathbb{E}_b(\xi)) \mathbb{E}_b(d \log x)$. As expected, only relative ξ 's and relative x 's are identified but not the overall levels $\mathbb{E}_b(\xi)$ and $\mathbb{E}_b(d \log x)$, since these are not pinned down by the underlying preference relation.

and the change in welfare is

$$EV^m \approx \Delta \log I - \mathbb{E}_{b_{t_0}}(\Delta \log p) - \frac{1}{2}(1 - \theta_0)Var_{b_{t_0}}(\Delta \log p) \quad (16)$$

$$- Cov_{b_{t_0}}(\Delta \log x, \Delta \log p) - \left[\Delta \log I - \mathbb{E}_{b_{t_0}}(\Delta \log p) \right] Cov_{b_{t_0}}(\varepsilon, \Delta \log p).$$

Consider the expenditure-switching terms in expression (15) one-by-one. If goods are substitutes, $\theta_0 > 1$, then welfare is convex in prices and variance in price changes boosts welfare by allow substitution towards goods that become relatively cheap. The second line of (15) captures the effect of taste shocks and income effects. If the composition of demand shifts in favor of goods that become relatively cheap, either due to taste shocks $Cov_b(\Delta \log x, \Delta \log p) < 0$ or income effects $Cov_b(\varepsilon, \Delta \log p)(\Delta \log I - \mathbb{E}_b(\Delta \log p)) < 0$, then real consumption increases.

Having understood expression (15), consider now expenditure-switching terms in expression (16). Welfare places the same weight on substitution effects as real consumption does. However, it places a larger weight on expenditure-switching due to income effects and taste shocks. Whereas $\Delta \log Y$ only takes into account expenditure-switching as it occurs over time, EV^m accounts for expenditure-switching due to income effects and taste shocks from the start. Therefore, expenditure-switching due to income and tastes are multiplied by 1/2 for $\Delta \log Y$ and 1 for EV^m . This implies that, for example, the increase in welfare from a price reduction in a good i with increasing demand (due to an increase in x_i or a relatively high ε_i) is not fully reflected in real consumption, implying $EV^m > \Delta \log Y$. Even if preferences are unstable or non-homothetic, real consumption strays from welfare only if income effects or taste shocks covary with price changes.¹⁶

2.4 Alternative Welfare Measures

In this section, we discuss two alternative welfare measures that are widely used in the literature: changes in utility indices and compensating variation.

¹⁶In Section 2.2, we pointed out that starting at b_{t_1} , computing welfare does not require knowledge of income elasticities or taste shocks if we know the elasticities of substitution. However, the approximation in (16) depends on income elasticities and taste shocks. The reason is because this approximation is around initial budget shares b_{t_0} . If we start with budget shares at t_1 and take an approximation as t_0 moves back in time, we get

$$EV^m \approx \Delta \log I - \mathbb{E}_b(\Delta \log p) + \frac{1}{2}(1 - \theta_0)Var_b(\Delta \log p),$$

where $\mathbb{E}_b(\cdot)$ and $Var_b(\cdot)$ are evaluated using budget shares at t_1 . Hence, starting at the terminal budget shares, EV^m only depends on substitution effects as in Example 1. Both expressions are valid second-order approximations and in either case, real consumption undercounts expenditure-switching caused by income effects or taste shocks.

Utility Index. A common approach in the literature is to use u_t as a measure of welfare and define the associated "ideal" price index $P_t \equiv e(p_t, u_t; x_t)/u_t$ by analogy to the homothetic and stable case. This is particularly common with CES preferences with either taste shocks or non-homotheticities (e.g. Comin et al., 2021; Redding and Weinstein, 2020; Ehrlich et al., 2019; Argente et al., 2021). If the utility function is not changing, then u_t obviously ranks choices according to the underlying preference relation, but the magnitude of the changes in u_t and the associated price index P_t are not interpretable. This is because they depend on the overall level of utility elasticities ζ , which are arbitrary. Furthermore, when preferences are non-homothetic, there is no normalization of the ζ parameters such that changes in u_t are equal to EV^m (see Online Appendix D.2).

This issue is even more severe when there are taste shocks. In this case, changes in u_t and P_t depend on the overall change in taste shifters x_t . Once again, the overall level of x_t in (6) is not pinned down by choice behavior. Therefore, to compute u_t one needs to assume a value for the overall level of x_t 's period-by-period. This means u_t no longer corresponds to the ranking of choices according to any preference relation, but is determined by the untestable assumption that determines the overall level of x_t 's in each period. For example, Redding and Weinstein (2020) set the average of x_t 's to be constant over time. However, the same data are consistent with the average of x_t 's rising or falling in arbitrary ways through time. Each assumption results in a different answer. As Muellbauer (1975) forcefully argues: "It is not possible to set up a market experiment which would unambiguously indicate whether the consumer's efficiency as a utility machine has changed." As he points out, "no such problems attend an ordinal cost-of-living index" like the ones we characterize.¹⁷

Compensating variation. Our baseline measure of welfare changes is equivalent variation. An alternative is the compensating variation, which is the reduction in t_1 income that makes the consumer indifferent to their t_0 budget set. As mentioned earlier, EV^m is a money-metric that ranks all choices, whereas CV^m is not. To see this, consider a household in t_0 choosing between (p_{t_1}, I_{t_1}) and (p'_{t_1}, I'_{t_1}) . It follows that $u'_{t_1} = v(p'_{t_1}, I'_{t_1}; x) > v(p_{t_1}, I_{t_1}; x) = u_{t_1}$ if, and only if, $e(p_{t_0}, u'_{t_1}; x)/I_{t_0} > e(p_{t_0}, u_{t_1}; x)/I_{t_0}$. That is, the household prefers whichever choice has the highest EV^m . Hence, EV^m is itself an index of utility. On the other hand, CV^m cannot be used to compare t_1 and t'_1 because CV^m for t_1 and t'_1 are expressed in terms of income adjustments in t_1 and t'_1 prices respectively. Compensating variation can be used to compare t_0 to t_1 and t_0 to t'_1 but it cannot be used to compare t_1 to t'_1 .

¹⁷See also Balk (1989) and Martin (2020) for related discussion of the cardinal and ordinal approaches.

Nevertheless, CV^m can be characterized using similar methods, as shown in Online Appendix A. For example, every result in the paper can be translated into compensating variation under initial preferences simply by reversing the flow of time. That is,

$$CV^m = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} b_i(p_t, u_{t_0}, x_{t_0}) \frac{d \log p_{it}}{dt} dt. \quad (17)$$

In particular, EV^m preserves the shape of the indifference curve at t_1 , whereas CV^m preserves the shape of the indifference curve at t_0 .¹⁸ To a second order starting at t_0 , for a consumer with non-homothetic CES preferences,

$$CV^m \approx \Delta \log I - \mathbb{E}_{b_{t_0}} (\Delta \log p) - \frac{1}{2} (1 - \theta_0) \text{Var}_{b_{t_0}} (\Delta \log p).$$

Comparing this with (16) reveals that EV^m accounts for expenditure-switching due to income effects whereas CV^m does not. Furthermore, up to a second-order approximation, $CV^m - \Delta \log Y = \Delta \log Y - EV^m$. Hence, the average of EV at final preferences and CV at initial preferences is equal to real consumption $\Delta \log Y$, up to a second order. See Online Appendix A for a generalization beyond non-homothetic CES.

3 Macroeconomic Changes in Welfare and Consumption

In the previous section we showed how to value changes in budget sets for a consumer who takes prices as given. For a whole society however, prices are endogenous to collective choices. In this section, we extend our analysis to show how to assign value to different production possibility frontiers (PPFs) rather than budget sets. The *macro*, or society-level, notion of welfare we define is not the same as micro welfare evaluated in general equilibrium.

We first introduce a closed, representative-agent, neoclassical economy with non-homothetic and unstable preferences. We generalize our definitions of welfare, now at the macroeconomic level, and present exact and approximate expressions for real GDP and welfare in terms of endogenous sufficient statistics. In the next section, Section 4, we solve for these endogenous sufficient statistics in terms of observable primitives. Without non-homotheticities and taste shocks, our economy is the same as the one studied by Baqaee and Farhi (2019). We extend their approximation formulas to capture real GDP and welfare

¹⁸Hence, calculating CV^m requires knowledge of initial budget shares and elasticities of substitution, whereas equivalent variation EV^m requires knowledge of final budget shares and elasticities of substitution. This means that EV^m at final preferences is more convenient for ex-post comparisons and CV^m at initial preferences is more convenient for counterfactuals.

when there are income effects and taste shocks.

3.1 Environment and Definitions

Consider a perfectly competitive neoclassical closed economy with a representative agent.¹⁹ Each good $i \in N$ has a production function

$$y_i = A_i G_i \left(\{m_{ij}\}_{j \in N}, \{l_{if}\}_{f \in F} \right),$$

where m_{ij} are intermediate inputs used by i and produced by j , and l_{if} denotes primary factor inputs used by i for each factor $f \in F$. The exogenous scalar A_i is a Hicks-neutral productivity shifter.²⁰ Without loss of generality, we assume that G_i has constant returns to scale because decreasing returns to scale can be captured by adding producer-specific factors. Furthermore A_i is Hicks-neutral without loss of generality. This is because we can capture non-neutral (biased) productivity shocks to input j for producer i by introducing a fictitious producer that buys from j and sells to i with a linear technology. A Hicks-neutral shock to this fictitious producer is equivalent to a non-neutral technology shock to i .

Let A be the $N \times 1$ vector of technology shifters and L be the $F \times 1$ vector of primary (exogenously given) factor endowments.²¹ The production possibility set (and its associated frontier) is the set of feasible consumption bundles that can be attained given A and L . Given our assumption that production functions have constant returns to scale, the PPF is linear if there is only one factor of production.

For each A , L , and x , we denote equilibrium prices (for both goods and factors) by $p(A, L, x)$ and aggregate income, or nominal GDP, by $I(A, L, x)$. These equilibrium prices and incomes are unique up to the choice of a numeraire.

Define the *macro indirect utility* function as the solution to the following planning problem:

$$V(A, L; x) = \max_c \{U(c; x) : c \text{ is feasible}\}.$$

This is the maximum amount of utility the economy can deliver given technologies (A, L)

¹⁹In Baqaee and Burstein (2021) we show that our results can be generalized to economies with heterogeneous agents and distortions. To measure the change in welfare from t_0 to t_1 , we ask: “*what is the minimum amount endowments in t_0 must change so that it is possible to make every consumer indifferent between t_0 and t_1 ?*” We show that the results in this paper generalize to economies without representative agents if we define social welfare in this way, and this definition collapses to the definition we use here when there is a representative agent. We capture distortions by including wedges as part of the primitives of the economy.

²⁰Note that unlike utility, physical output of each individual good has interpretable units and can be measured directly, without the need to use a money-metric.

²¹Allowing for endogenous labor-leisure choice requires including the time endowment in L and leisure in the consumption bundle c .

and preferences \succeq_x . Whereas the micro indirect utility takes prices as given and lets consumers pick any point in their budget set (even if such a point is not feasible at the economy-wide level), the macro indirect utility function takes the PPF as the primitive and lets society only pick feasible points in the production possibility set. The first welfare theorem implies that the competitive equilibrium decentralizes the planning problem above with prices determined in equilibrium.

Consider shifts in the PPF as technologies and factor endowments change from (A_{t_0}, L_{t_0}) to (A_{t_1}, L_{t_1}) , along with changes in preferences from x_{t_0} to x_{t_1} . We generalize our microeconomic measure of welfare in the following way.

Definition 5 (Macro Welfare). The change in welfare measured using the *macro equivalent variation* with *final preferences* is $EV^M(A_{t_0}, L_{t_0}, A_{t_1}, L_{t_1}; x_{t_1}) = \phi$ where ϕ solves

$$V(A_{t_1}, L_{t_1}; x_{t_1}) = V(A_{t_0}, e^\phi L_{t_0}; x_{t_1}).$$

The superscript M in EV^M represents the fact that this is the *macro* equivalent variation, in contrast to EV^m for the *micro* welfare measure. In words, EV^M is the proportional change in initial factor endowments required to make the consumer with preferences $\succeq_{x_{t_1}}$ indifferent between the PPF defined by $(A_{t_0}, e^\phi L_{t_0})$ and the new PPF, defined by (A_{t_1}, L_{t_1}) .

Intuitively, EV^M expresses utility changes in terms of factor endowments. That is, EV^M is itself an index of utility and $(A, L) \succeq_x (A', L')$ if, and only if, $EV^M(A_{t_0}, L_{t_0}, A, L; x) \geq EV^M(A_{t_0}, L_{t_0}, A', L', x)$. The macro equivalent variation, EV^M , is a useful metric because it ranks PPFs without reference to endogenous prices.

Definition 5 shows that EV^M is a generalization of EV^m in the sense that the two coincide when the PPF is the same as the budget constraint. However, EV^M is also a generalization of *consumption-equivalents* commonly used to measure welfare in macroeconomics (e.g. Lucas 1987). To see this, note that EV^M can also be defined in the following way.

Definition 6 (Alternative Definition of Macro Welfare). Let \mathcal{C}_t denote the production possibility set associated with (A_t, L_t) . Then $EV^M = \phi$, where ϕ solves

$$V(\mathcal{C}_{t_1}; x_{t_1}) = V(e^\phi \mathcal{C}_{t_0}; x_{t_1}),$$

and $e^\phi \mathcal{C}_{t_0} \equiv \{e^\phi c : c \in \mathcal{C}_{t_0}\}$.

In words, EV^M also measures the proportional shift in \mathcal{C}_{t_0} necessary to make $\succeq_{x_{t_1}}$ indifferent between $e^\phi \mathcal{C}_{t_0}$ and \mathcal{C}_{t_1} . This definition is isomorphic to Definition 5. This is because scaling all factor endowments by a constant shifts out the PPF by the same constant under our assumptions.

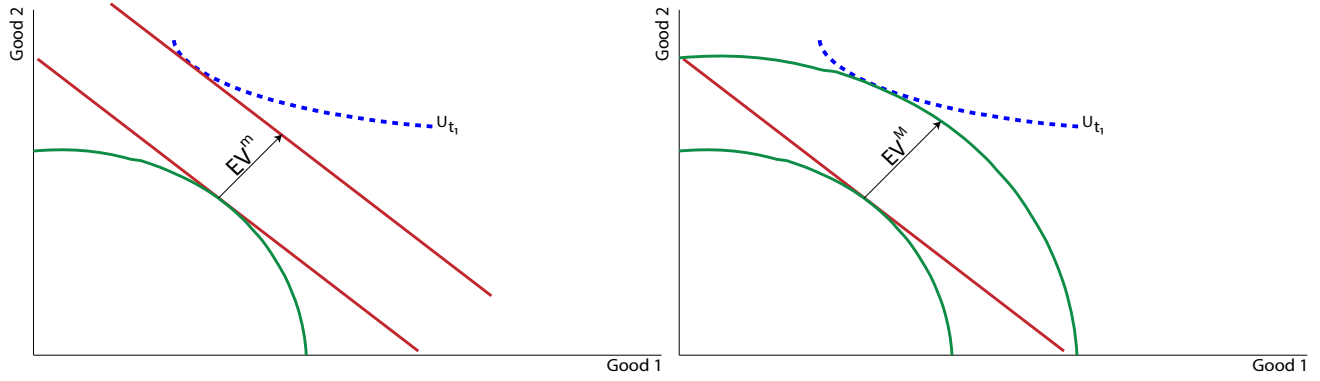


Figure 1: The left panel shows the change in t_0 income that makes the household with final preferences indifferent between the budget constraint in t_0 and t_1 . The right panel shows the change in t_0 endowments that makes the household with final preferences indifferent between the PPF in t_0 and t_1 . EV^m and EV^M are equal if the PPF is linear, or if preferences are homothetic and stable.

For a graphical representation of EV^m and EV^M , see Figure 1. Macro and micro welfare answer different questions. To see the difference, consider a situation where households age between t_0 and t_1 but technologies and factor endowments stay the same. Because the PPF is unchanged, the change in macro welfare is zero by construction. However, if the PPF is nonlinear, the relative price of goods changes between t_0 and t_1 : prices rise for goods that become more desirable. In this case, EV^m necessarily falls even though the PPF is unchanged. Hence, EV^m does not rank PPFs for a society.

The following examples further illustrate why micro welfare can be a misleading measure of the welfare impact of technological change when preferences are non-homothetic or unstable.

Example 2 (Micro versus macro welfare and growth). Consider economies that are endowed with cows, and suppose each cow produces one unit of steak and one unit of offal. At t_0 the economy is endowed with one cow. We consider what happens if we double the number of cows at t_1 . For any stable and homothetic preferences, macro and micro welfare doubles ($EV^M = EV^m = \log 2$).

Panel (a) in Figure 2 considers a case with non-homothetic but stable preferences. For this example, suppose that the consumer at t_0 considers steak and offal to be perfect substitutes because the consumer is poor and only cares about total calories. Hence, the relative price of offal and steak is one and the consumption bundle is one unit of steak and one unit offal in t_0 . At t_1 , the consumer consumes two units of steak and two units of offal. If steak is an extreme luxury good, then at t_1 , the relative price of steak is much higher than offal. Macro welfare measures the distance between the original PPF and the new PPF, which is twice as big since the economy in t_1 produces double the number of both steaks and offal. However, in this example, micro welfare only rises by an arbitrarily small number ε . That

is, even though the economy is twice as good, micro welfare has increased by almost zero. Intuitively, in the initial equilibrium, steak was relatively cheap (because consumers were poor), so initial income in t_0 only needs to be raised by ϵ to allow the consumer to reach the t_1 indifference curve.

In other words, micro welfare measures the fact that an infinitesimal consumer with an endowment of $1 + \epsilon$ cows could purchase a total of $2 + \epsilon$ steaks. However, if all consumers' endowments were raised by ϵ , and all consumers tried to sell their offal to buy steak, then the relative price of steak would quickly rise because the aggregate supply of steaks has only risen by ϵ . This means that in practice, consumers with $1 + \epsilon$ cows are not as well off as they thought they would be. Hence, to make them indifferent, we would have to give them more endowment — doing this would change relative prices again. Macro welfare is the fixed point to this problem, where the increase in the endowment is exactly enough to make consumers indifferent taking into account the fact that relative prices are endogenous.²²

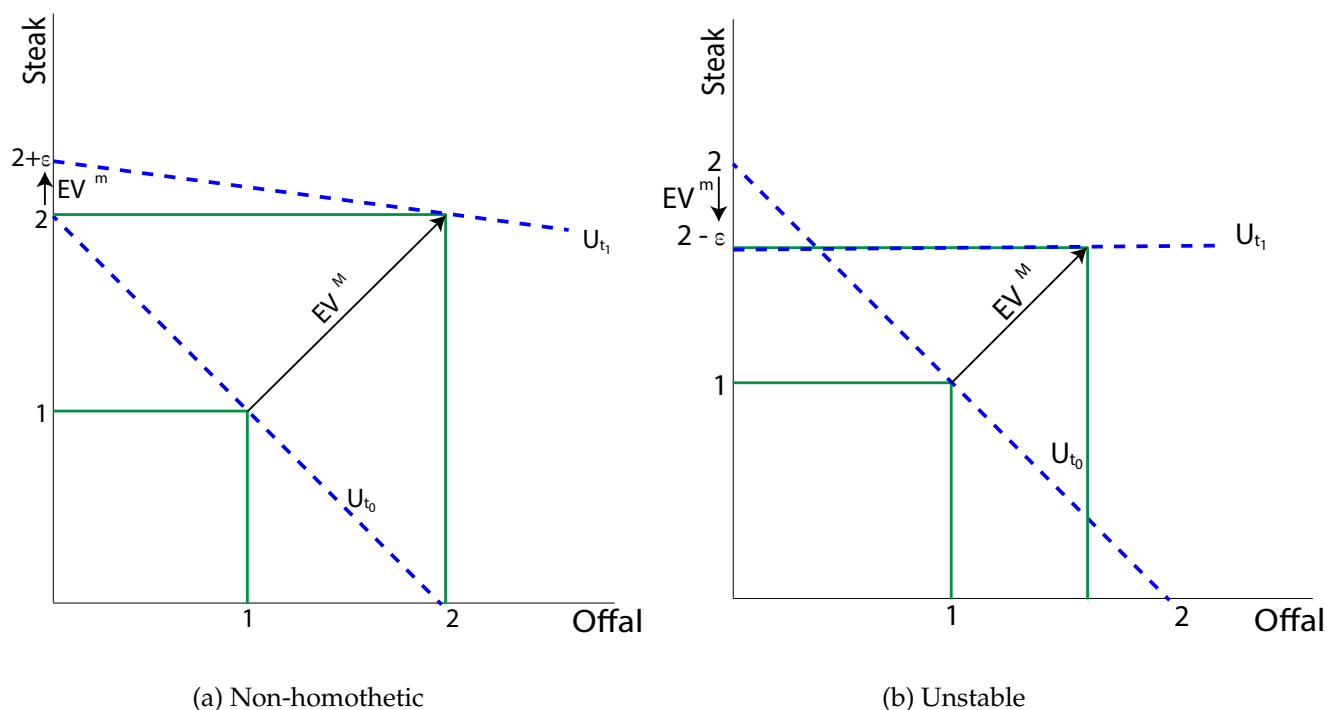


Figure 2: The left panel shows an example with non-homothetic preferences in which the PPF doubles but micro welfare increases by an arbitrarily small ϵ . The right panel shows an example in which the PPF expands by $2 - \epsilon$ but micro welfare falls. Indifference curves are dashed blue lines, PPFs are solid green lines.

Panel (b) in Figure 2 considers a case with homothetic but unstable preferences instead.

²²Surprisingly, even though macro welfare is the solution to this seemingly complicated fixed point problem, our analytical results will show that it is in fact simpler to characterize than micro welfare in general equilibrium.

Starting with the same PPF as before in t_0 , suppose that in t_1 the number of cows increases by a factor $2 - \varepsilon$. Due to a concurrent advertising campaign, consumer preferences shift so that in t_1 steak becomes more popular than offal (the indifference curves in Figure 2 (b) become flatter). In this example, since the number of cows has increased by a factor $2 - \varepsilon$, the macro welfare change is $\log(2 - \varepsilon)$. However, micro welfare declines because the consumer with t_1 preferences strictly prefers the t_0 budget set. Intuitively, a consumer with t_1 preferences could sell the offal in t_0 and purchase a total of 2 steaks, which is preferred to $2 - \varepsilon$ steaks she can get in t_1 . That is, even though the economy has almost doubled in terms of its productive capabilities, micro welfare in general equilibrium has fallen. Hence, micro welfare in general equilibrium is not a good measure of how changes in choice sets affect societal welfare.²³

The previous examples illustrate that using the initial budget set to represent the initial PPF is deceptive, since the initial budget set reflects both the technologies and demand in t_0 . Our macroeconomic notion of welfare accounts for the endogenous changes in prices by comparing the initial and final PPFs rather than the initial and final budget sets.

When relative prices do not respond to consumers' choices (i.e. the PPF is linear), then for a given primitive shock, macro and micro welfare are always the same. Alternatively, if preferences are homothetic and stable, then macro and micro welfare are the same, regardless of the shape of the PPF. The following proposition formalizes this:

Proposition 4 (Macro vs. Micro Welfare). *Consider changes in technologies A , factor quantities L , and tastes x . Macro and micro welfare changes are equal ($EV^m = EV^M$) if preferences are stable and homothetic, or if the PPF is linear.*

We provide a quantitative illustration of the difference between micro and macro welfare using the COVID-19 crisis in Section 5.

3.2 Characterization of Welfare and Real GDP

To characterize macro welfare, we define a Hicksian (or compensated) economy indexed by (A, L, u, x) , where (A, L) determines the PPF and (u, x) specifies an indifference curve corresponding to $U(c; x) = u$.

Definition 7 (Hicksian economy). Consider a fictitious consumer whose preferences are represented by the expenditure function $\tilde{e}(p, \tilde{u}) \equiv e(p, u; x)\tilde{u}$, for fixed values of u and

²³ Although EV^m and EV^M can have different signs when preferences are unstable, they have the same sign when preferences are stable. This is because u_t is equal in the micro and macro problem (by the first welfare theorem), and $EV^m > 0$ and $EV^M > 0$ if, and only if, $u_{t_1} > u_{t_0}$.

x . In words, $\tilde{e}(p, \tilde{u})$ maps prices p and utility values \tilde{u} into required expenditures of the fictitious consumer. The *Hicksian* economy, indexed by (A, L, u, x) , is an economy with PPF (A, L) and expenditure function $\tilde{e}(p, \tilde{u})$.

In the definition above, u and x are parameters that define the fictional consumer's preferences. When the true consumer's preferences, represented by the expenditure function $e(\cdot)$, are stable and homothetic, then they coincide with the fictional consumer's preferences, represented by $\tilde{e}(\cdot)$. In this case, the Hicksian economy is the same as the uncompensated original economy. We refer to variables, like prices and sales, in the Hicksian economy as Hicksian variables and index them by (A, L, u, x) . We refer to variables in the uncompensated original economy as *observed* variables and index them by t .

Denote the sales shares relative to GDP of each good or factor i by

$$\lambda_i = \frac{p_i y_i}{I} \mathbf{1}(i \in N) + \frac{p_i L_i}{I} \mathbf{1}(i \in F),$$

where $\mathbf{1}$ is an indicator function, and N and F are the set of goods and factors. The sales share λ_i is often referred to as the *Domar* weight of i . Note that referring to λ_i as a “share” is an abuse of language since $\sum_{i \in N} \lambda_i > 1$ whenever there are intermediate inputs.

To characterize macro welfare and compare it to real GDP, we use *Hicksian* sales shares, $\lambda(A, L, u, x)$, which are the sales shares in a Hicksian economy. The Hicksian sale shares are the general equilibrium counterpart to partial equilibrium Hicksian budget shares. The following proposition shows that changes in macro welfare is the integral of Hicksian sales shares with respect to changes in technologies and factor quantities.

Proposition 5 (Macro Welfare). *For any smooth path of technologies and factor quantities that unfold as a function of time t , changes in macro welfare are*

$$EV^M = \underbrace{\int_{t_0}^{t_1} \sum_{i \in N} \lambda_i(A_t, L_t, u_{t_1}, x_{t_1}) \frac{d \log A_{it}}{dt} dt}_{\text{welfare-relevant TFP growth}} + \underbrace{\int_{t_0}^{t_1} \sum_{i \in F} \lambda_i(A_t, L_t, u_{t_1}, x_{t_1}) \frac{d \log L_{it}}{dt} dt}_{\text{welfare-relevant factor growth}}. \quad (18)$$

In words, growth accounting for welfare should be based on compensated or Hicksian sales shares evaluated at current technology but final preferences and utility. We define the first N summands of (18) to be changes in *welfare-relevant TFP* and the last F summands are changes in welfare due to changes in factor inputs. As we show below, this is analogous to how measured TFP is defined using real GDP.

To apply Proposition 5, the only thing we need to know about preferences is how $e(p, u_{t_1}; x_{t_1})$ varies as a function of prices for given u_{t_1} and x_{t_1} . This uniquely pins down the fictional consumers' preferences that generate the Hicksian shares $\lambda(A_t, L_t, u_{t_1}, x_{t_1})$. In

other words, we need to know the budget shares and elasticities of substitution at the final allocation but not income elasticities or taste shocks.²⁴

We now consider changes in real GDP, defined using the Divisia index for final goods as $\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} b_{it} d \log c_{it}$. The following result shows that changes in real GDP can be calculated by integrating observed sales shares with respect to technology changes.

Proposition 6 (Real GDP). *Given a path of technologies, factor quantities, and tastes that unfold as a function of time t , the change in real GDP is*

$$\Delta \log Y = \underbrace{\int_{t_0}^{t_1} \sum_{i \in N} \lambda_i(A_t, L_t, u_t, x_t) \frac{d \log A_{it}}{dt} dt}_{TFP \text{ growth}} + \underbrace{\int_{t_0}^{t_1} \sum_{i \in F} \lambda_i(A_t, L_t, u_t, x_t) \frac{d \log L_{it}}{dt} dt}_{\text{factor growth}}, \quad (19)$$

where $\lambda(A_t, L_t, u_t, x_t)$ are observed sales shares λ_{it} at t .

In (19), the first N summands are equal to measured TFP, and the last F summands are the growth in real GDP caused by changes in factor endowments. Proposition 6 is a slight generalization of Hulten (1978) to environments with unstable and non-homothetic final demand. In general equilibrium, the sales shares play the role that budget shares played in partial equilibrium. Whereas in partial equilibrium, integrating budget shares with respect to prices yielded real consumption, in general equilibrium integrating sales shares with respect to technologies and factors yields real GDP. If the PPF is not moving, that is, technology A and endowments L are constant, then pure preference shocks have no effect on real GDP even though they can change allocations.

Propositions 5 and 6 show that the difference between macro welfare and real GDP can be understood in terms of the distinction between Hicksian and observed sales shares. If preferences are homothetic and stable, then $\lambda(A, L, u_{t_1}, x_{t_1}) = \lambda_{it}$ and the change in welfare is equal to real GDP.

To get more intuition for the difference between welfare and real GDP, we use a second-order approximation of the response of real GDP and welfare to technology and preference shocks as $t_1 - t_0 = \Delta t \rightarrow 0$. To make the formulas more compact and without loss of generality, when we write local approximations we abstract from shocks to factor endowments ($\Delta \log L = 0$).

Proposition 7 (Approximate Real GDP and Macro Welfare). *Up to a second order approxima-*

²⁴Following the observation made in Section 2.4, for compensating variation at initial preferences, we need to know elasticities of substitution at the initial allocation instead of the final one.

tion, the change in real GDP is

$$\Delta \log Y \approx \sum_{i \in N} \lambda_i \Delta \log A_i + \frac{1}{2} \sum_{i \in N} \Delta \lambda_i \Delta \log A_i, \quad (20)$$

and the change in welfare is

$$EV^M \approx \sum_{i \in N} \lambda_i \Delta \log A_i + \sum_{i \in N} \Delta \lambda_i \Delta \log A_i - \underbrace{\frac{1}{2} \sum_{j \in N} \Delta \log A_j \frac{\partial \lambda_i}{\partial \log A_j} \Delta \log A_i}_{\text{expenditure-switching due to substitution effects}}, \quad (21)$$

where $\partial \lambda / \partial \log A$ is the partial derivative of the Hicksian sales share with respect to technology, and all terms are evaluated at t_0 .

We discuss equations (20) and (21) in turn. The first term in (20) is the Hulten-Domar formula. The second term in (20), emphasized by Baqaee and Farhi (2019), captures nonlinearities due to changes in sales shares. Intuitively, if sales shares decrease for those goods with higher productivity growth, then real GDP growth slows down.

Equation (21) shows that the gap between macro welfare and real GDP is similar to that for our micro results (the signs are flipped because a positive productivity shock reduces prices). Specifically, real GDP takes into consideration changes in sales shares *along* the equilibrium path. These changes in sales shares could be induced by technology shocks but they could also be due to changes in preferences and non-homotheticities. Equation (21) shows that welfare puts more weight on changes in sales shares that are due to demand shocks and non-homotheticities than those due to substitution effects. That is, real GDP “undercorrects” for changes in shares caused by non-homotheticities or changes in preferences, similar to the partial equilibrium counterparts in Proposition 2 and 3.

4 Structural Macro Results and Analytic Examples

The results in Section 3 are reduced-form in the sense that they take changes in observed and compensated sales shares as given. In this section, we solve for changes in these endogenous objects in terms of microeconomic sufficient statistics. For clarity, we restrict attention to nested-CES economies. The general case is in Online Appendix I, and the intuition is very similar. This section also contains some worked-out analytical examples.

Nested-CES economies. Household preferences are non-homothetic CES, with expenditure function (6), where once again θ_0 is the elasticity of substitution across consumption

goods. Production uses nested-CES aggregators. We adopt the following representation. Each good $i \in N$ is produced with the production function

$$y_i = A_i G_i \left(\{m_{ij}\}_{j \in N}, \{l_{if}\}_{f \in F} \right) = A_i \left(\sum_{j \in N} \omega_{ij} m_{ij}^{\frac{\theta_i-1}{\theta_i}} + \sum_{f \in F} \omega_{if} l_{if}^{\frac{\theta_i-1}{\theta_i}} \right)^{\frac{\theta_i}{\theta_i-1}},$$

where ω_{ij} and ω_{if} are constant parameters, and θ_i is the elasticity of substitution across inputs for good i . Any nested-CES production network with an arbitrary pattern of nests and elasticities can be represented in this way if we treat each CES aggregator as a separate producer (see Baqaee and Farhi 2019).

Input-output matrix. We stack the expenditure shares of the household, all producers, and all factors into the $(1 + N + F) \times (1 + N + F)$ input-output matrix Ω . The first row corresponds to the household. To highlight the special role played by the household, we index it by 0, which means that the first row of Ω is equal to the household's budget shares introduced above ($\Omega_0 = b'$, with $b_i = 0$ for $i \notin N$). The next N rows correspond to the expenditure shares of each producer on every other producer and factor. The last F rows correspond to the expenditure shares of the primary factors (which are all zeros, since primary factors do not require any inputs).

Leontief inverse matrix. The Leontief inverse matrix is the $(1 + N + F) \times (1 + N + F)$ matrix defined as

$$\Psi \equiv (Id - \Omega)^{-1} = Id + \Omega + \Omega^2 + \dots,$$

where Id is the identity matrix. The Leontief inverse matrix $\Psi \geq Id$ records the *direct and indirect* exposures through the supply chains in the production network. We partition Ψ in the following way:

$$\Psi = \left[\begin{array}{c|ccc|ccc} 1 & \lambda_1 & \cdots & \lambda_N & \lambda_{N+1} & \cdots & \lambda_{N+F} \\ 0 & \Psi_{11} & \cdots & \Psi_{1N} & \Psi_{1N+1} & \cdots & \Psi_{1N+F} \\ 0 & & \ddots & & & & \\ 0 & \Psi_{N1} & & \Psi_{NN} & \Psi_{NN+1} & \cdots & \Psi_{NN+F} \\ \hline 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & 1 & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{array} \right].$$

The first row and column correspond to final demand (good 0). The first row is equal to the vector of sales shares for goods and factors λ' . The next N rows and columns correspond to goods, and the last F rows and columns correspond to the factors. The right F columns of Ψ represent the network-adjusted factor intensities of each good. The sum of network-adjusted factor intensities for every i is equal to one, $\sum_{f \in F} \Psi_{if} = 1$ because the factor content of every good is equal to one. In our results below we use the identity that $\lambda' = b' \Psi$.

4.1 General characterization for nested-CES economies

According to Propositions 5 and 6, changes in real GDP and welfare can be computed by integrating observed and Hicksian sales shares with respect to technology shocks. The following proposition pins down sales shares as a function of primitives. This proposition can then be used in combination with Propositions 5 and 6 to calculate exact changes in real GDP and welfare. For readability, we again assume away shocks to factor endowments.

Proposition 8 (Characterization for nested-CES economies). *At any point in time t , changes in observed prices and the Leontief inverse are pinned down by the following differential equations:*

$$d \log p_{it} = - \underbrace{\sum_j \Psi_{ijt} d \log A_{jt}}_{\text{upstream TFP changes}} + \underbrace{\sum_{f \in F} \Psi_{ift} d \log \lambda_{ft}}_{\text{upstream factor price changes}}, \quad (22)$$

$$d\Psi_{ilt} = \underbrace{\sum_j \Psi_{ijt} (\theta_j - 1) \text{Cov}_{\Omega_{(j,:),t}} \left(-d \log p_t, \Psi_{(:,l),t} \right)}_{\text{substitution effect}} + \underbrace{\mathbf{1}_{\{i=0\}} \text{Cov}_{\Omega_{(0,:),t}} \left(d \log \mathbf{x}_t + \varepsilon_t d \log Y_t, \Psi_{(:,l),t} \right)}_{\text{taste shocks and income effect}}, \quad (23)$$

where changes in sales shares are $d\lambda_{it} = d\Psi_{0it}$ and changes in real GDP are $d \log Y_t = \sum_i \lambda_{it} d \log A_{it}$. As before, the demand shifter for good i is $d \log x_{it} = d \log x_{it} - (\xi_i / \mathbb{E}_{b_t}[\xi]) \mathbb{E}_{b_t}[d \log x]$ and the income elasticity of demand for good i is $\varepsilon_{it} = 1 + (1 - \theta_0) (\xi_i / \mathbb{E}_{b_t}[\xi] - 1)$.²⁵

To compute changes in real GDP using Proposition 6, we need to know $\lambda_{it} = \lambda(A_t, L_t, u_t, x_t)$ for $t \in [t_0, t_1]$. These are solutions to the differential equations in (22) and (23) given some boundary condition that pins down Ψ at some point in time $t \in [t_0, t_1]$.

On the other hand, to compute macro welfare, we need Hicksian sales shares $\lambda(A_t, L_t, u_{t_1}, x_{t_1})$ as a function of t . These are solutions to the same differential equations, except that the

²⁵ For all of these expressions, the summations are evaluated over all goods and factors, so that i and $j \in \{0\} + N + F$, $\text{Cov}_{\Omega_{(j,:),t}}(\cdot)$ is the covariance using the j th row of Ω at time t as the probability weights, and $\Psi_{(:,l),t}$ is the l th column of the Leontief inverse at time t .

terms involving taste shocks and income effects in equation (23) are set to zero. In this case, the boundary condition is that the Leontief inverse at t_1 is equal to the observed Leontief inverse Ψ_{t_1} at t_1 . Therefore, if Ψ_{t_1} is observed, we can calculate Hicksian sales shares between t_0 and t_1 by starting (23) at t_1 and going backwards to t_0 . This process does not require knowledge of either the income elasticities ε nor the taste shocks $\Delta \log x$.

Each term in the differential equations in Proposition 8 has a clear interpretation. Equation (22) captures the fact that the price of each good $d \log p_i$ is determined by its (direct and indirect) exposure to the price of inputs j and factors f (captured by Ψ_{ijt} and Ψ_{ift} at time t).

On the other hand, the equation for $d \log \Psi_{ilt}$ shows that changes in the Leontief inverse are determined by substitutions by j , if j is an intermediary between i and l , as well as income and substitution effects if i is the household ($i = 0$). Finally, the Hicksian version of this equation is identical except that it does not adjust expenditures due to income effects and taste shocks along the transition.

Remark 1 (Micro Welfare). Proposition 8 can also be used to compute changes in microeconomic welfare EV^m in a general equilibrium model. To do this, we compute $p(A_{t_0}, L_{t_0}, u_{t_0}, x_{t_0})$ and $p(A_{t_1}, L_{t_1}, u_{t_1}, x_{t_1})$ using Proposition 8 and then plug these price changes into equation (9). Unlike macroeconomic welfare, microeconomic welfare changes generically require knowledge of both income elasticities and taste shocks because they are needed to calculate the initial and or final prices.

To build more intuition about the difference between real GDP and macro welfare, consider economies with only a single factor of production. In this case the system is simplified since we know that the economy's single primary factor always has a sales share equal to one. This allows for a simple characterization of both welfare and real GDP up to a second-order approximation.

Proposition 9 (Approximate Macro Welfare vs GDP: Single Factor). *When the economy has one factor of production, up to a second order approximation, the change in real GDP is*

$$\begin{aligned} \Delta \log Y \approx & \sum_{i \in N} \lambda_i \Delta \log A_i + \frac{1}{2} \sum_j \lambda_j (\theta_j - 1) \text{Var}_{\Omega_{(j,:)}} \left(\sum_{i \in N} \Psi_{(:,i)} \Delta \log A_i \right) \\ & + \frac{1}{2} \text{Cov}_{\Omega_{(0,:)}} \left(\Delta \log x + \left(\sum_{i \in N} \lambda_i \Delta \log A_i \right) \varepsilon, \sum_{i \in N} \Psi_{(:,i)} \Delta \log A_i \right), \end{aligned} \quad (24)$$

and welfare is

$$EV^M \approx \Delta \log Y + \frac{1}{2} \text{Cov}_{\Omega_{(0,:)}} \left(\Delta \log \mathbf{x} + \left(\sum_{i \in N} \lambda_i \Delta \log A_i \right) \varepsilon, \sum_{i \in N} \Psi_{(:,i)} \Delta \log A_i \right), \quad (25)$$

where λ , Ω , Ψ , ε , and $\Delta \log \mathbf{x}$ are evaluated at t_0 .

Proposition 9 is a general equilibrium counterpart to Proposition 3. We first discuss real GDP. The first term in Equation (24) is the Hulten-Domar term. The other terms are second-order terms resulting from the fact that sales shares change in response to shocks. The first term captures nonlinearities due to the fact that sales shares respond to changes in relative prices caused by technology shocks (these effects are emphasized by Baqaee and Farhi, 2019). The terms on the second line of (24), which are the ones we focus on in this paper, capture changes in sales shares due to changes in tastes or non-homotheticities.

Equation (25) shows that while real GDP correctly accounts for substitution due to supply shocks, in order to measure welfare it needs to be corrected for expenditure-switching due to changes in final demand caused by taste shocks or income effects. Whereas in partial equilibrium, the gap between welfare and real GDP is proportional to the covariance of price and demand shocks (see Proposition 3), equation (25) shows that in general equilibrium, the relevant statistic is the covariance of demand shocks with a network-adjusted notion of supply shocks.

4.2 Analytical Examples

We now work through some simple examples to illustrate the forces that drive a gap between real GDP and welfare.

Example 3 (Correlated Supply and Demand Shocks). We start with the simplest possible example, a one sector model without any intermediates. In this case, sales shares are just budget shares $\lambda_i = b_i = \Omega_{0i}$. Therefore, Proposition 9 simplifies to

$$EV^M - \Delta \log Y \approx \frac{1}{2} (\text{Cov}_b(\Delta \log \mathbf{x}, \Delta \log A) + \text{Cov}_b(\varepsilon, \Delta \log A) \mathbb{E}_b[\Delta \log A]). \quad (26)$$

Welfare changes are greater than the change in real GDP if productivity and demand shocks (i.e., shifts in demand curves) are positively correlated. This could happen either because preferences exogenously change to favor high productivity growth goods, $\text{Cov}_b(\Delta \log \mathbf{x}, \Delta \log A) > 0$, or income effects favor high productivity growth goods, $\text{Cov}_b(\varepsilon, \Delta \log A) \mathbb{E}_b[\Delta \log A] > 0$. When shifts in demand are orthogonal to shifts in sup-

ply, to a second-order approximation, real GDP measures welfare correctly. By Proposition 4, in this example, macro welfare EV^M is the same as micro welfare EV^m .

We work through some simple examples with multiple factors of production to illustrate how nonlinear PPFs affect the previous results.

Example 4 (Decreasing Returns to Scale). Consider the one-sector model without intermediate inputs in Example 3 but now suppose that production functions are non-constant-returns-to-scale. Specifically, the production for good i is

$$y_i = A_i L_i^\gamma,$$

where L_i is labor and γ need not equal one. To apply our propositions to this economy, where producers have non-constant-returns production functions, we introduce a set of producer-specific factors in inelastic supply, and suppose that each producer has a Cobb-Douglas production function that combines a common factor with elasticity γ and a producer-specific factor with elasticity $1 - \gamma$. This means that our economy has $1 + N$ factors.

For simplicity, suppose that preferences are homothetic ($\varepsilon_i = 1$ for every i) but potentially unstable ($\Delta \log x \neq 0$). We apply Proposition 7 to compute the difference between welfare and real GDP. To do this, we first use Proposition 8 to compute changes in observed and Hicksian sales shares due to demand shocks and then plug this into Proposition 7 to get the difference between welfare and real GDP up to a second order approximation:

$$EV^M - \Delta \log Y \approx \frac{1}{2} \frac{\text{Cov}_b(\Delta \log x, \Delta \log A)}{1 + (\theta_0 - 1)(1 - \gamma)}. \quad (27)$$

This expression simplifies when we have constant-returns to scale ($\gamma = 1$) or when the elasticity of substitution is unity ($\theta_0 = 1$). Outside of these cases, complementarities ($\theta_0 < 1$) amplify the effect of preference shocks under decreasing returns to scale ($\gamma < 1$). Intuitively, if preferences shift in favor of some good, the price of that good rises due to decreasing returns to scale. The fact that the price of the good increases raises the sales share of that good due to complementarities, which creates a feedback loop, raising prices of the good further, and causing additional substitution. In other words, in the decreasing returns to scale model with complementarities, sales shares respond more strongly to demand shocks. Given that sales shares respond more strongly to demand shocks, the necessary adjustment to correct real GDP is larger.

Example 5 (Macro vs. Micro Welfare Change). Finally, we demonstrate the difference between macro and micro welfare changes using the previous example. The economy in

the previous example has multiple factors and unstable preferences. Therefore, macro and micro notions of welfare are different because the PPF is no longer linear.

To illustrate this difference, suppose that only preference shocks are active (there are no supply shocks $\Delta \log A = 0$ and $\Delta \log L = 0$). Since the PPF is being held constant, macro-welfare changes are also zero. Micro-welfare changes, on the other hand, are not equal to zero. Specifically, by Proposition 3, micro welfare improves $EV^m > 0$ if preference shocks negatively covary with price changes. Using Proposition 8, changes in prices are

$$d \log p_i = \frac{(1 - \gamma)}{(1 + (\theta_0 - 1)(1 - \gamma))} (d \log x_i - \mathbb{E}_b[d \log \mathbf{x}]).$$

If there are decreasing returns, $\gamma < 1$, then a positive demand shock for i raises the price of i . This initial change in the price is amplified by general equilibrium forces if goods are complements and mitigated if goods are substitutes (this is the multiplier in the denominator). We can now apply Proposition 3 to obtain micro welfare, up to a second order,

$$EV^m \approx -\frac{1}{2} \frac{(1 - \gamma)}{(1 + (\theta_0 - 1)(1 - \gamma))} \text{Var}_b(\Delta \log \mathbf{x}) \neq 0 = EV^M.$$

With decreasing returns to scale ($\gamma < 1$), micro welfare is negative because the demand shock increases the prices of goods the consumer now values more. From a micro perspective, where the agent takes the budget sets as given, the agent is worse off. Of course, from a societal perspective, welfare has not changed, since the production possibility set of the economy has not changed.

Online Appendix E contains additional examples showing how input-output connections can amplify or mitigate the gap between macro welfare and real GDP.

5 Applications

In this section we consider three applications of our results. The first application is about long-run growth, quantifying the difference between welfare-relevant and measured aggregate productivity growth in the presence of income effects and demand instability. The second application is about short-run fluctuations and shows that correlated product-level supply and demand shocks within industries drive a gap between measured real consumption and welfare even in the short-run. Our final application is a business cycle event study, where we use the COVID-19 recession to demonstrate the difference between macroeconomic and microeconomic welfare using a quantitative model. We also use this model to show that demand instability can make real GDP an unreliable metric for changes in

production.

5.1 Application I: Long-Run Growth and Structural Transformation

As economies grow, sectors with low productivity growth tend to expand compared to sectors with faster productivity growth. This means that over time, aggregate productivity growth is increasingly determined by those sectors whose productivity growth is slowest. This phenomenon is often called Baumol's cost disease.

Following Nordhaus (2008), aggregate productivity growth between t_0 and t_1 can be decomposed into two terms:

$$\Delta \log TFP = \sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it_0} \Delta \log A_{it} + \underbrace{\sum_{t=t_0}^{t_1} \sum_{i \in N} (\lambda_{it} - \lambda_{it_0}) \Delta \log A_{it}}_{\text{Baumol Adjustment}},$$

where λ_{it} is the sales shares of industry i in period t and $\Delta \log A_{it}$ is the growth in gross-output productivity over each time period.²⁶ The first term captures changes in aggregate TFP if industrial structure had remained fixed, and the second term is the adjustment attributed to the fact that sales shares change over time. The second-term captures the importance of Baumol's cost disease.

Proposition 5 implies that, for the purposes of welfare, changes in sales shares due to income effects or demand instability must be treated differently to changes in sales shares due to substitution effects. In particular, for EV^M , the relevant measure of the change in TFP is

$$\Delta \log TFP^{ev} = \sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it_0} \Delta \log A_{it} + \underbrace{\sum_{t=t_0}^{t_1} \sum_{i \in N} (\lambda_{it} - \lambda_{it_0}) \Delta \log A_{it}}_{\text{Baumol Adjustment}} + \underbrace{\sum_{t=t_0}^{t_1} \sum_{i \in N} (\lambda_{it}^{ev} - \lambda_{it}) \Delta \log A_{it}}_{\text{Welfare Adjustment}},$$

where $\lambda^{ev} \equiv \lambda(A_t, L_t, u_{t_1}, x_{t_1})$ is the Hicksian sales shares of each industry.

Two polar extremes. Computing the Hicksian sales shares to obtain $\Delta \log TFP^{ev}$ requires an explicit structural model of the economy. However, there are two polar cases in which $\Delta \log TFP^{ev}$ can be calculated without specifying the detailed model. The first extreme

²⁶Technically, this is an approximation, since we define aggregate TFP in continuous time but the data are measured in discrete time (at annual frequency). However, this approximation error, resulting from the fact that the Riemann sum is not exactly equal to the integral is likely to be negligible in practice. At our level of disaggregation, long run TFP growth is very similar if we weight sectors using sales shares at time t or time t and $t + 1$ averages.

is when demand is stable and homothetic, so that changes in sales shares are due only to relative price changes (substitution effects). The second extreme is when there are no substitution effects in sales shares (as in a Cobb-Douglas economy), and changes in sales shares are only due to income effects or demand instability. If structural transformation is driven by a combination of substitution effects and non-homotheticities or demand instability, then the change in welfare TFP will be somewhere in between these two cases, as discussed in Online Appendix F. The following corollary of Proposition 5 summarizes the change in welfare-TFP in these polar cases.

Corollary 1. *If changes in sales shares are due only due only to substitution effects, then*

$$\Delta \log TFP^{ev} = \Delta \log TFP = \sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it} \Delta \log A_{it}.$$

If changes in sales shares are due only to non-homotheticity or instability of demand, then

$$\Delta \log TFP^{ev} = \Delta \log TFP + \sum_{t=t_0}^{t_1} \sum_{i \in N} (\lambda_{it_1} - \lambda_{it}) \Delta \log A_{it} = \sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it_1} \Delta \log A_{it}.$$

In the first case, since preferences are homothetic and stable, $\lambda(A_t, L_t, u_{t_1}, x_{t_1}) = \lambda_{it}$, so welfare-TFP is equal to TFP in the data. In the second case, since there are no substitution effects in production or demand, compensated sales shares do not respond to productivity changes, so $\lambda(A_t, L_t, u_{t_1}, x_{t_1}) = \lambda_{it_1}$.

To quantify Corollary 1, we use US-KLEMS data on sales shares and TFP growth for 61 private-sector industries between 1947 and 2014. We calculate changes in industry-level gross-output TFP following the methodology of Jorgenson et al. (2005) and Carvalho and Gabaix (2013).²⁷ Figure 3 plots EV^M comparing 1947 to subsequent years under alternative assumptions about substitution and income elasticities. For comparisons that are close to 1947, the change in welfare is not very sensitive to our assumptions about elasticities. This is because at higher frequencies, the shocks are small and industry sales shares are reasonably stable. However, the assumptions about substitution and income elasticities play a larger role over longer horizons. Comparing 1947 to 2014, the constant-initial-sales-share term grows by around 58 log points (or 78%), whereas the chain-linked change in aggregate TFP grew by around 47 log points (60%). Hence, Baumol's cost-disease caused aggregate TFP to fall by 10 log points and reduced aggregate productivity growth by around 23%.

If we assume that structural transformation is due solely to income effects and taste

²⁷For each industry, the change in TFP is itself a chain-weighted index calculated as output growth minus share-weighted input growth. Inputs are industry-level measures of materials, labor, and capital services.

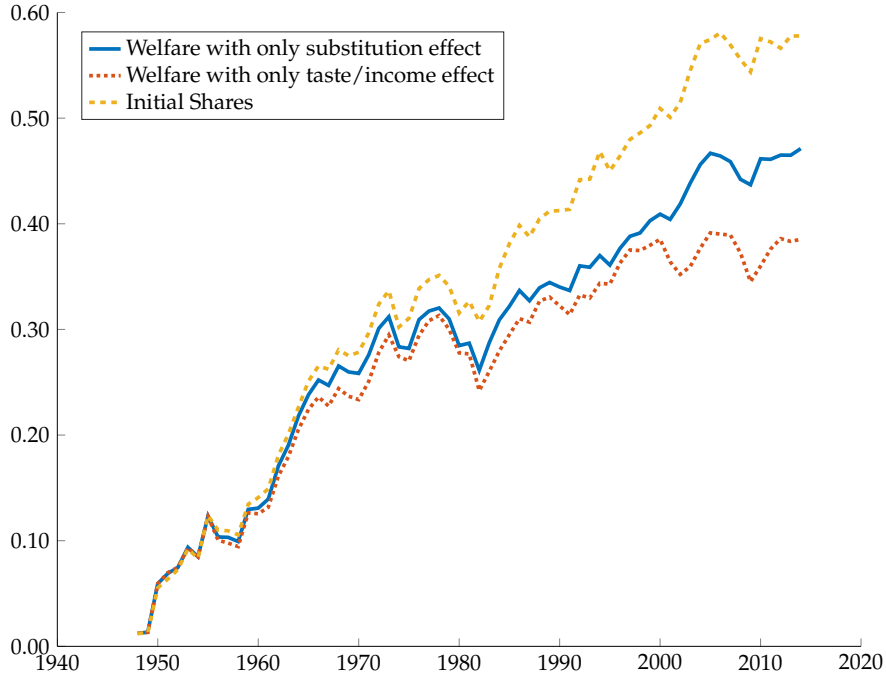


Figure 3: Growth in welfare-relevant TFP (in logs) from 1947 to 2014 using US-KLEMS. The dashed yellow line uses initial 1947 shares to calculate TFP changes. The solid blue and dotted red lines measure the increase in welfare-relevant TFP between 1947 and each year t between 1948 and 2014 under alternative assumptions about income and substitution elasticities. The blue line assumes that sales shares change only due to substitution effects (welfare-relevant TFP is equal to measured chained-aggregate TFP). The red line assumes that sales shares change only due to income effects (or demand instability).

shocks, then by Corollary 1 the growth in welfare-relevant TFP from 1947-2014 was 37 log points (46%) instead of the measured 47 log points (60%) — that is, to say, a 23% additional reduction in the growth rate.

Intuitively, welfare-based productivity increases less than TFP because, relative to 1947, preferences in 2014 favor low-productivity growth sectors such as services (due to either income effects or demand instability). These sectors were cheaper compared to manufacturing in 1947 than in 2014. This means that at 1947 prices, when services were relatively cheap, households require less income growth to be indifferent between their budget constraint in 1947 and the one in 2014.²⁸

²⁸This intuition is flipped for compensating variation at 1947 preferences. Preferences in 1947 favor manufacturing over services. Therefore, at 2014 prices, when manufacturing is relatively cheap, households require a larger reduction in 2014 income to make them indifferent to 1947. More generally, if structural transformation is purely due to income effects or preference instability, then welfare-based productivity growth using CV at initial preferences is given by initial sale-share weighted productivity growth, $\sum_{t=t_0}^{t_1} \sum_{i \in N} \lambda_{it_0} \Delta \log A_{it}$ (which corresponds to the Initial Shares line in Figure 3), so for this exercise, the Baumol adjustment is not welfare-relevant. The fact that EV at final preferences and CV at initial preferences are different stems from the fact that they answer different questions, so EV uses demand from 2014 whereas CV uses demand from 1947.

In Online Appendix F, we provide some quantitative illustrations away from the two polar extremes discussed above. In this appendix, we compute welfare changes for different values of elasticities of substitution in consumption and production using Proposition 8. We show that welfare-relevant TFP is closer to measured TFP if the elasticity of substitution across industries (in consumption or production) is significantly lower than one.

5.2 Application II: Inflation with Product-Level Taste Shocks

In the previous application, we considered a long-run industry-level application. With industry-level data, the gaps between welfare and chain-weighted indices are usually modest over the short-run because industry-level sales shares are relatively stable at high frequency. However, this does not mean that these issues are absent in short-run data.

Whereas industry sales shares are stable at high frequency, firm or product-level sales shares are highly volatile even over the very short-run. If firms' or products' supply and demand shocks are correlated, then measured industry-level output is biased relative to what is relevant for welfare.²⁹ We provide an empirical illustration of the magnitude of the biases caused by taste shocks in product-level data using the Nielsen Consumer Panel database. In the body of the paper, we only briefly describe the dataset, and refer readers to Online Appendix H for more details. The Nielsen Consumer Panel tracks the purchasing behavior of about 40,000 to 60,000 panelists every year from 2004 to 2019 for a wide variety of non-durable consumer goods (food, non-food groceries, general merchandise, etc.). A product in the data is defined by its unique Universal Product Code (UPC), and each product is assigned to a module. Our balanced sample covers roughly 820 modules. Panelists in the sample are assigned weights, allowing purchases by the panel to be projected to a nationally representative sample.

We model national demand for UPCs in a given module using a homothetic but unstable CES functional form. We set $t_1 = 2019$ and then for each year $t_0 < 2019$, we calculate welfare-relevant deflators for t_0 and t_1 preferences module-by-module for continuing products using equation (9). The price of each UPC in each year is calculated as the ratio of national expenditures on that UPC over units sold over the whole year. For each t_0 , we include only UPCs purchased in each quarter of each year between t_0 and t_1 . In other words, we abstract from product entry and exit by focusing on the continuing-goods price

²⁹In Online Appendix G, we formally show that the biases in industry-level data are not diversified away as we aggregate, even if all products are infinitesimal in their industry and all industries are infinitesimal in the aggregate economy. Furthermore, we provide conditions under which the within-industry biases are, to a second-order, linearly separable from the across-industry biases. That is, the overall bias is the sum of the cross-industry bias (that we studied in the previous section) plus additional biases driven by within-industry covariance of supply and demand shocks.

index (see Section 6 for how to deal with product entry-exit when preferences are unstable). To combine module-level inflation rates into a single number, we assume preferences in both t_0 and t_1 across modules are Cobb-Douglas. For the same set of UPCs, we also compute the change in inflation as measured by a chained Tornqvist index (a discrete time approximation to the Divisia index).

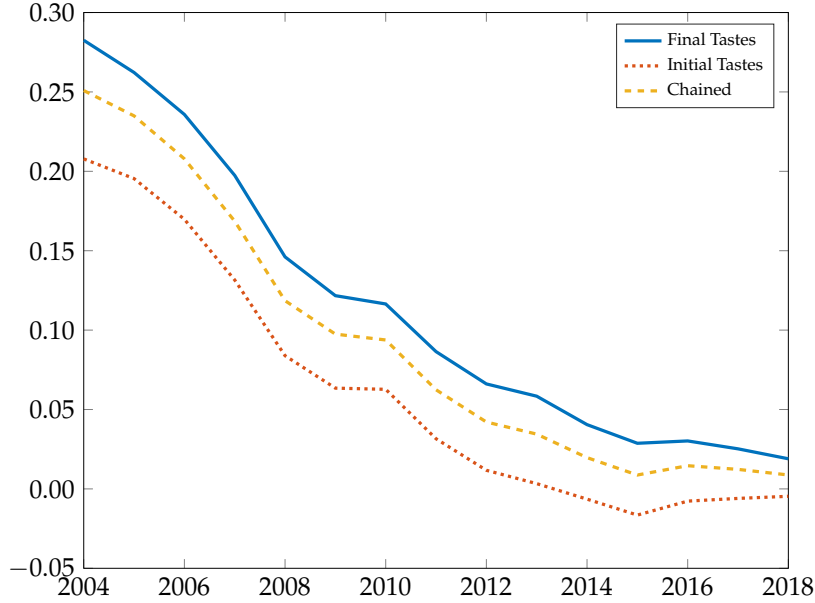


Figure 4: Welfare-relevant and chain-weighted inflation rates for continuing products. The welfare-relevant rates are computed assuming that the elasticity of substitution across UPCs in the same module is 4.5 and the elasticity of substitution across modules is one.

Figure 4 displays the welfare-relevant and chained inflation rates for each t_0 assuming that the elasticity of substitution across UPCs in the same module is 4.5.³⁰ Starting in 2018, inflation for 2019 preferences is around 1 percentage point higher than the chained index. Intuitively, this is because changes in prices and changes in demand shocks between t_0 and t_1 are positively correlated. That is, goods that are more popular in 2019 had relatively higher inflation rates. Following the logic of Proposition 3, this means that the chained index understates inflation for final preferences. The gap increases as we go farther back in time because the shocks are persistent and cumulate. Over the whole sample, the gap widens to 4.3 percentage points.

On the other hand, for initial preferences, the chained index overstates the inflation rate. This is because goods that are relatively more popular in t_0 compared with t_1 (i.e.

³⁰Estimating the elasticity of substitution is beyond the scope of this paper, therefore, for our empirical illustration we draw on estimates from the literature. An elasticity of 4.5 is at the lower range of estimates reported by Redding and Weinstein (2020) and Jaravel (2019). In Online Appendix H, we report results for higher and lower elasticities. We find that the size of the bias is increasing in the elasticity of substitution. In this sense, the results in Figure 4 are relatively conservative.

goods that became less popular over time) experience lower inflation rates. Since goods that become more popular are also the goods that have higher inflation rates, it is natural that inflation with final tastes is higher than inflation with initial tastes. The two series are different because they answer different questions.^{31,32}

We report robustness with respect to the elasticity of substitution parameter in Online Appendix H. In this appendix, we show that the size of the bias gets smaller as θ gets closer to one. This is because in the data changes in prices and changes in expenditure shares are approximately uncorrelated. When demand is Cobb-Douglas, changes in expenditure shares are driven only by taste shocks, and so taste shocks are roughly uncorrelated with price changes. Hence, following the logic of Proposition 3, the bias is smaller in the Cobb-Douglas case and larger if the calibrated elasticity is greater than 4.5.

5.3 Application III: the COVID-19 Recession

Our final application examines how real GDP, microeconomic welfare, and macroeconomic welfare were affected during the COVID-19 recession. The COVID-19 recession is an interesting case study since sectoral expenditure shares changed substantially during this time, these changes were not explainable via changes in observed prices alone, and the movements in demand curves were correlated with movements in supply curves. These are exactly the conditions under which micro welfare, macro welfare, and real GDP can diverge from each other.³³

In this application, we do not attempt to measure the welfare costs of COVID-19 itself. This is because households do not make choices over whether or not they live in a world with COVID-19. Therefore, their preferences about COVID-19 itself are not revealed by

³¹If we were to average initial and final tastes for each t_0 , then we would get a series that is fairly close to the chained index in this example. Since the chained index is, to a second-order, equal to the average of the series with initial and final tastes, this suggests that the second-order approximation is relatively accurate for this example. In Online Appendix H, we also report results using monthly, rather than annual, data. In that case, the chained measure is not as close to the average of initial and final tastes because the second-order approximation is less accurate.

³²The “taste-adjusted inflation rate” in Redding and Weinstein (2020) is lower than the chain-weighted inflation. This may appear related to our finding that inflation for initial tastes is lower than chained inflation. However, for initial tastes, our welfare-relevant inflation is simply exact hat-algebra with initial expenditure shares, discarding any information on subsequent taste shocks. This is different to Redding and Weinstein (2020) whose measure does account for changing tastes, but models an increase in tastes as equivalent to a reduction in price and pins down the overall level of taste shifters by assuming that their average stays constant through time. For a better understanding of why our approach is fundamentally different to that of Redding and Weinstein (2020), see Online Appendix C.

³³Cavallo (2020) argues that, during this episode, the fact that price indices were not being chained at high enough frequency led to “biases” in official measures of inflation. However, as we have argued, chaining is only theoretically valid if expenditure-switching is caused by substitution effects, and not if expenditure-switching is caused by shocks to demand.

their choices. Instead, we ask a more modest question: how does the household value changes in prices (micro welfare) and changes in production (macro welfare), holding *fixed* the presence of COVID-19.

To study this episode, we use a modified version of the quantitative model introduced in Section 4. Since this is a short-run application, we assume that factor markets are segmented by industry, so that labor and capital in each industry is inelastically supplied. We calibrate share parameters to match the 71 industry US input-output table in 2018 (we exclude government sectors) from the Bureau of Economic Analysis, and consider a range of elasticities of substitution. Following Baqaee and Farhi (2022), we model the COVID-19 recession as a combination of negative sectoral employment shocks and sectoral taste shifters. We hit the economy with a vector of primitive supply and demand shocks. The reductions in sectoral employment are calibrated to match peak-to-trough reductions in hours worked by sector from January 2020 to May 2020. The primitive demand shifters in household demand are calibrated to match the observed peak-to-trough change in personal consumption expenditures by sector from January 2020 to May 2020 (conditional on the supply shocks and the elasticities of substitution).³⁴

We consider three different calibrations informed by empirical estimates from Atalay (2017) and Boehm et al. (2019): high complementarities, medium complementarities, and no complementarities (Cobb-Douglas). The high complementarity scenario sets the elasticity of substitution across consumption goods to be 0.7, the one across intermediates to be 0.01, across value-added and materials to be 0.3, and the one between labor and capital to be 0.2. The medium complementarities case sets the elasticity of substitution across consumption goods to be 0.95, the one across intermediates to be 0.01, across value-added and materials to be 0.5, and the one between labor and capital to be 0.5. The Cobb-Douglas calibration sets all elasticities of substitution equal to unity.

Table 1 displays welfare changes between January 2020 and May 2020 in the calibrated model. We report separately micro and macro welfare based on pre-COVID ($t_0 = \text{Q1-2018}$) and COVID ($t_1 = \text{Q2-2020}$) preferences. Recall that micro and macro welfare are not equal in this economy because the PPF is nonlinear.

Table 1 shows that micro welfare falls by more than macro welfare under COVID preferences. This is because, as shown in Example 5, when demand rises for goods in the presence of decreasing returns to scale, micro welfare falls more than macro welfare. Intuitively, if household demand for, say, toilet paper rises, then this raises the price of toilet

³⁴Changes in labor by sector and personal consumption expenditures, used to calibrate supply and demand shocks, are taken from Baqaee and Farhi (2022). For related analysis of COVID-19 induced supply shocks, see e.g. Bonadio et al. (2021) and Barrot et al. (2021). For related analysis of COVID-19 induced demand shocks, see Cakmakli et al. (2020).

Table 1: Change in micro and macro welfare with pre-COVID and COVID preferences given the supply and demand shocks between February 2020 to May 2020. Chained real GDP is computed assuming supply and demand shocks arrive simultaneously.

Elasticities	High compl.	Medium compl.	Cobb-Douglas
Micro pre-COVID preferences	-11.7%	-9.1%	-8.7%
Micro COVID preferences	-13.2%	-12.3%	-10.9%
Macro pre-COVID preferences	-16.2%	-12.5%	-10.8%
Macro COVID preferences	-10.1%	-9.4%	-9.0%
Chained real GDP	-12.1%	-10.6%	-9.8%

paper and reduces micro welfare compared to macro welfare. That is, a single consumer with COVID preferences, who values toilet paper and takes prices as given, is made worse off by the fact that demand for toilet paper rose. However, society as a whole does not take prices as given. Therefore, when comparing the t_0 PPF to the t_1 PPF, the fact that toilet paper was cheaper in t_0 due to lower demand is irrelevant.

This pattern is reversed for pre-COVID preferences. Macro welfare at pre-COVID preferences fell by a lot because supply contracted in sectors that were popular pre-COVID (e.g., airline travel). Micro welfare losses at pre-COVID preferences are smaller because demand also fell in those sectors (e.g. airline travel) and this reduction in demand reduced their equilibrium prices. Hence, the consumer with pre-COVID preferences, who still likes airline travel, is not as worse off as the society that still likes airline travel.

This illustrates that micro and macro welfare answer different questions, and the answers to these questions can be quantitatively very different. Furthermore, comparing columns of Table 1 shows that the magnitude of these differences depend on the details of the production structure like the extent of complementarities in production. As we raise the elasticities of substitution in production closer to unity (Cobb-Douglas), the differences between macro and micro notions become less dramatic. This is because the PPF becomes less curved.

In Table 1, we also compute real GDP assuming supply and demand shocks arrive simultaneously and linearly over time. Under this assumption, real GDP is somewhere in between the welfare measures at t_0 and t_1 preferences for both macro and micro. However, this is only because of our assumption that the supply and demand shocks arrive simultaneously and at the same speed. If we change the path of supply and demand shocks, real GDP changes value (even though the initial and final allocation are not changing). For example, if the supply shocks arrive before the demand shocks, then real GDP equals macro

welfare changes at t_0 preferences. On the other hand, if demand shocks arrive before the supply shocks, then real GDP equals macro welfare changes at t_1 preferences.

Hence, if the supply and demand shocks do not disappear in exactly the same way as they arrived, measured real GDP after the recovery can be higher or lower than it was before the crisis, even if the economy returns exactly to its pre-COVID allocation. That is, if in the downturn demand shocks arrive before supply shocks (so real GDP falls by roughly 10% in the high complementarities case, according to Table 1) and, in the recovery, demand shocks disappear before the supply shocks (so real GDP rises by roughly 16%), then real GDP is as much as 6% higher when comparing pre-shock real GDP to post-recovery real GDP. This is despite the fact that every price and quantity is the same when comparing the pre-shock allocation to the post-recovery allocation. Hence, during episodes where final demand is unstable, chained real GDP and consumption are unreliable guides for measuring output or welfare, even if we chain in continuous time.³⁵

6 Extensions

In this section, we briefly summarize how our theoretical results can be used in dynamic settings and extended to account for new goods.

Dynamic Economies

At an abstract level, our results can be applied to dynamic economies by using the Arrow-Debreu formalism of indexing goods by period of time and state of nature (as in Basu et al., 2022). In a dynamic economy the utility function is intertemporal and capital accumulation must be treated as an intertemporal intermediate good, as advocated by Barro (2021). In Online Appendix J, we specialize our results further to show how Proposition 5 can be used to make welfare comparisons in a dynamic multi-sector model with production of consumption goods and investment goods, with unstable and non-homothetic preferences. We show that steady-state to steady-state macro welfare is equal to the change in nominal consumption expenditures deflated by the exact-algebra CES price index associated with the t_1 indifference curve, exactly as for the partial equilibrium microeconomic welfare in

³⁵This is related to a problem known as “*chain drift*” bias in national accounting. Chain drift occurs when a chained index registers an overall change between t_0 and t_1 even though all prices and quantities in t_0 and t_1 are identical. This is a specific manifestation of path dependence of chained indices (see Hulten, 1973) and, by the gradient theorem for line integrals, it must be driven by either demand instability, income effects, or approximation errors due to discreteness. Chain drift bias can appear when movements in prices and quantities are oscillatory, where changes that take place over some periods are reversed in subsequent periods. Welfare changes do not exhibit chain drift since, by definition, they depend only on t_0 and t_1 variables.

expression (9), despite the fact that the dynamic general equilibrium economy has infinitely many factors.

Extensive margin.

In our discussion, we do not explicitly deal with the new goods problem (i.e., goods discontinuously appearing or disappearing). The classic treatment assumes that consumers start and stop consuming goods because the price falls from or goes to infinity. Using stable CES preferences, Feenstra (1994) provides a simple-to-implement formula to adjust price indices for entry and exit of goods. However, when preferences are unstable, it is possible that consumers start or stop consuming a good not because of a change in its availability but because of a change in their tastes. In this section, using the commonly used CES specification, we show that the welfare implications of this can be profound.

Consider a consumer whose preferences are homothetic CES with taste shifters x and elasticity of substitution $\theta_0 > 1$. Good i is *unavailable* if its price is infinite, $p_i = \infty$, and *available* otherwise. Demand for good i may be zero either because the good is unavailable ($p_i = \infty$) or because the consumer does not value the good ($x_i = 0$).

Split the set of goods that consumers value at t_1 into three sets: (1) \mathcal{C} : continuing goods consumed in both periods, $b_{it_0} > 0$ and $b_{it_1} > 0$; (2) \mathcal{N} : newly consumed goods that either were unavailable at t_0 ($p_{it_0} = \infty$ and $p_{it_1} < \infty$) or that were available ($p_{it_0} < \infty$) but are valued at t_1 and not at t_0 ($x_{it_0} = 0$ and $x_{it_1} > 0$); (3) \mathcal{X} : exiting goods that become unavailable at t_1 ($p_{it_0} < \infty$ and $p_{it_1} = \infty$).

The following proposition derives the change in welfare.

Proposition 10 (Welfare with Product Entry/Exit). *The change in equivalent variation at t_1 tastes is given by*

$$EV^m = \log \frac{I_{t_1}}{I_{t_0}} - \log \frac{P_{t_1}}{P_{t_0}}, \quad (28)$$

where

$$\frac{P_{t_1}}{P_{t_0}} = \left(\frac{b_{t_1}^c \left(\frac{P_{t_1}^c}{P_{t_0}^c} \right)^{\theta_0-1} + (1 - b_{t_1}^c) \left(\frac{P_{t_1}^n}{P_{t_0}^n} \right)^{\theta_0-1}}{1 - b_{t_0}^{x_{t_1}}} \right)^{\frac{1}{\theta_0-1}}. \quad (29)$$

In this expression $b_{t_1}^c$ is the expenditure share on continuing goods in t_1 , and $P_{t_1}^c / P_{t_0}^c$ and $P_{t_1}^n / P_{t_0}^n$ is the change in a CES price index for continuing, \mathcal{C} , and new, \mathcal{N} , goods under t_1 tastes. Finally, $b_{t_0}^{x_{t_1}}$ is the expenditure share on exiting goods under t_1 tastes and p_{t_0} prices.

Applying equation (29) requires three pieces of information:

- i. The t_1 share of continuing goods, $b_{t_1}^c$, and changes in the t_1 price index for continuing goods $P_{t_1}^c / P_{t_0}^c$.
- ii. The price index for newly consumed goods $P_{t_1}^n / P_{t_0}^n$, which combines newly available goods, for which $p_{i_{t_1}} / p_{i_{t_0}} = 0$, and goods that were available in both periods but consumers did not have tastes for at t_0 , for which $p_{i_{t_1}} / p_{i_{t_0}} \neq 0$.
- iii. The counterfactual share of exiting goods at t_0 prices but t_1 preferences $b_{t_0}^{x_{t_1}}$. It is reasonable to think that $b_{t_0}^{x_{t_1}} \in [0, b_{t_0}^{x_{t_0}}]$. The lower bound takes the position that exiting goods are no longer valuable to the consumer with t_1 tastes, and the upper bound takes the position that exiting goods are not relatively more valuable to the consumer with t_1 tastes than the consumer with t_0 tastes (i.e., demand curves for those goods that exited did not shift out).

By making different assumptions about the bounds, we can derive three noteworthy special cases of expression (29).

No extensive margin. Assume that all goods are continuing ($b_{t_1}^c = 1$). Then the price deflator is the same as in expression (9):

$$\frac{P_{t_1}}{P_{t_0}} = \frac{P_{t_1}^c}{P_{t_0}^c} = \left(\sum_{i \in \mathcal{C}} b_{it_1} \left(\frac{p_{it_1}}{p_{it_0}} \right)^{\theta_0 - 1} \right)^{\frac{1}{\theta_0 - 1}}. \quad (30)$$

This is the assumption we make in Section 5.2, since we only compute the price index for continuing goods.

Feenstra (1994) with taste shocks. Suppose that newly consumed goods at t_1 were unavailable at t_0 ($P_{t_1}^n / P_{t_0}^n = 0$). Furthermore, suppose that, under t_1 tastes but t_0 prices, the share of expenditures on exiting goods equals the observed share of exiting goods at t_0 : ($b_{t_0}^{x_{t_1}} = b_{t_0}^{x_{t_0}} = 1 - b_{t_0}^c$). In this case, equation (29) reduces to

$$\frac{P_{t_1}}{P_{t_0}} = \left(\frac{b_{t_1}^c}{b_{t_0}^c} \right)^{\frac{1}{\theta_0 - 1}} \frac{P_{t_1}^c}{P_{t_0}^c}.$$

The term in front of the continuing price index coincides with the well-known new-goods adjustment with CES preferences due to Feenstra (1994). Note that P_{t_1} / P_{t_0} is still not the same as Feenstra (1994) because $P_{t_1}^c / P_{t_0}^c$ is calculated for fixed x_{t_1} tastes and hence is not the Sato-Vartia price index. In other words, $P_{t_1}^c / P_{t_0}^c$ is the price index in equation (30).

Entry/exit only due to taste shocks. Suppose that all goods were available in t_0 and t_1 , but some goods are consumed in t_1 and not t_0 because of changes in tastes. In this case, $b_{t_0}^{x_{t_1}} = 0$, so

$$\frac{P_{t_1}}{P_{t_0}} = \left(b_{t_1}^c \left(\frac{P_{t_1}^c}{P_{t_0}^c} \right)^{\theta_0-1} + (1 - b_{t_1}^c) \left(\frac{P_{t_1}^n}{P_{t_0}^n} \right)^{\theta_0-1} \right)^{\frac{1}{\theta_0-1}}. \quad (31)$$

In this case, the price index is a weighted CES aggregator of the price index for continuing goods and the price index for newly consumed goods. The price index increases less than the price index for continuing goods if inflation for newly consumed goods is less than inflation for continuing goods.³⁶

7 Conclusion

In this paper, we provide a toolkit for studying how welfare changes in response to changes in budget sets or production possibility sets allowing for preference instability and non-homotheticity. In contrast to the case of stable and homothetic preferences, there is no model-free statistic like chain-weighted real consumption that simultaneously provides answers to different welfare questions (i.e. equivalent or compensating variation, initial or final preferences, partial or general equilibrium). When preferences are non-homothetic and unstable, there are different welfare questions one may ask, and they have different answers. We characterize these measures and show the difference between them and chain-weighted real consumption.

Although our motivation and applications have focused on shocks across time, an interesting avenue to explore is welfare comparisons across locations (see e.g. Deaton, 2003, and Argente et al., 2021). Whereas in a temporal context, the preferences of today are typically more relevant than the preferences of yesterday, in a spatial context, both locations' preferences are equally interesting. The distinction between macroeconomic and microeconomic welfare is also important in a spatial context. Comparing budget constraints in one location to another may be misleading as a way to compare the technologies of two economies. This is because, even if PPFs in both locations are exactly the same, the relative price of goods households value more in one location will be lower in the other location.³⁷

³⁶For example, the model in Arkolakis (2016) predicts that changes in tastes induced by advertising will be correlated with changes in physical productivity, whereby more productive firms will expend more resources on advertising. In this case price of newly consumed goods increase on average less than continuing goods.

³⁷This implies that consumers from one location prefer the budget set in the other location, which resembles the Gerschenkron (1951) effect that the relative GDP of a country is higher when evaluated at another

References

- Acemoglu, D. and V. Guerrieri (2008). Capital deepening and nonbalanced economic growth. *Journal of political Economy* 116(3), 467–498.
- Aghion, P., A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li (2019). Missing growth from creative destruction. *American Economic Review* 109(8), 2795–2822.
- Alder, S., T. Boppart, and A. Muller (2022). A theory of structural change that can fit the data. *American Economic Journal: Macroeconomics* 14(2), 160–206.
- Argente, D. O., C.-T. Hsieh, and M. Lee (2021). Measuring the cost of living in mexico and the us. *American Economic Journal: Macroeconomics*. Forthcoming.
- Arkolakis, C. (2016). A unified theory of firm selection and growth. *The Quarterly Journal of Economics* 131(1), 89–155.
- Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal: Macroeconomics* 9(4), 254–80.
- Atkin, D., B. Faber, T. Fally, and M. Gonzalez-Navarro (2020, March). Measuring Welfare and Inequality with Incomplete Price Information. NBER Working Papers 26890, National Bureau of Economic Research, Inc.
- Auer, R., A. Burstein, S. M. Lein, and J. Vogel (2022, February). Unequal expenditure switching: Evidence from switzerland. Working Paper 29757, National Bureau of Economic Research.
- Balk, B. M. (1989). Changing consumer preferences and the cost-of-living index: theory and nonparametric expressions. *Journal of Economics* 50(2), 157–169.
- Baqae, D. and A. Burstein (2021). Aggregate welfare and output with heterogeneous agents.
- Baqae, D., A. Burstein, and Y. Koike-Mori (2022, October). A new method for measuring welfare with income effects using cross-sectional data. Working Paper 30549, National Bureau of Economic Research.
- Baqae, D. and E. Farhi (2022). Supply and demand in disaggregated keynesian economies with an application to the covid-19 crisis. *American Economic Review* 112(5), 1397–1436.
- Baqae, D. R. and E. Farhi (2019). The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten’s Theorem. *Econometrica* 87(4), 1155–1203.
- Barro, R. J. (2021). Double Counting of Investment. *The Economic Journal* 131(638), 2333–2356.
- Barrot, J.-N., B. Grassi, and J. Sauvagnat (2021). Sectoral effects of social distancing. In *AEA Papers and Proceedings*, Volume 111, pp. 277–81.
- Basu, S., L. Pascali, F. Schiantarelli, and L. Serven (2022). Productivity and the welfare of nations. *Journal of the European Economic Association* 20(4), 1647–1682.
- Baumol, W. J. (1967). Macroeconomics of unbalanced growth: the anatomy of urban crisis. *The American economic review*, 415–426.
- Boehm, C. E., A. Flaaen, and N. Pandalai-Nayar (2019). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 tōhoku earthquake. *Review of Economics and Statistics* 101(1), 60–75.
- Bonadio, B., Z. Huo, A. A. Levchenko, and N. Pandalai-Nayar (2021). Global supply chains in the pandemic. *Journal of international economics* 133, 103534.
- Boppart, T. (2014). Structural change and the kaldor facts in a growth model with relative price effects and non-gorman preferences. *Econometrica* 82(6), 2167–2196.
- Buera, F. J. and J. P. Kaboski (2009, 04-05). Can Traditional Theories of Structural Change Fit The Data? *Journal of the European Economic Association* 7(2-3), 469–477.

country’s prices (the grass is greener in the other side). However, while the Gerschenkron effect is a spatial version of the discrepancy between Laspeyres and Paasche indices, our result is driven by the endogeneity of relative prices to demand forces and not by substitution bias.

- Buera, F. J., J. P. Kaboski, and R. Rogerson (2015). Skill biased structural change. Technical report, National Bureau of Economic Research.
- Cakmakli, C., S. Demiralp, S. Kalemli-Ozcan, S. Yesiltas, and M. A. Yildirim (2020, May). Covid-19 and emerging markets: An epidemiological multi-sector model for a small open economy with an application to turkey. Working Paper 27191, National Bureau of Economic Research.
- Carvalho, V. and X. Gabaix (2013). The Great Diversification and its undoing. *The American Economic Review* 103(5), 1697–1727.
- Carvalho, V. M. and A. Tahbaz-Salehi (2019). Production networks: A primer. *Annual Review of Economics* 11, 635–663.
- Cavallo, A. (2020, June). Inflation with covid consumption baskets. Working Paper 27352, National Bureau of Economic Research.
- Caves, D. W., L. R. Christensen, and W. E. Diewert (1982). The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica*, 1393–1414.
- Comin, D., D. Lashkari, and M. Mestieri (2021). Structural change with long-run income and price effects. *Econometrica* 89(1), 311–374.
- Cravino, J., A. A. Levchenko, and M. Rojas (2019, September). Population Aging and Structural Transformation. NBER Working Papers 26327, National Bureau of Economic Research, Inc.
- Deaton, A. (2003). Adjusted indian poverty estimates for 1999-2000. *Economic and political Weekly*, 322–326.
- Deaton, A. and J. Muellbauer (1980). *Economics and consumer behavior*. Cambridge university press.
- Diewert, W. E. (1976). Exact and superlative index numbers. *Journal of econometrics* 4(2), 115–145.
- Domar, E. D. (1961). On the measurement of technological change. *The Economic Journal* 71(284), 709–729.
- Ehrlich, G., J. Haltiwanger, R. Jarmin, D. Johnson, and M. D. Shapiro (2019). Minding your ps and qs: Going from micro to macro in measuring prices and quantities. In *AEA Papers and Proceedings*, Volume 109, pp. 438–43.
- Fally, T. (2022). Generalized separability and integrability: Consumer demand with a price aggregator. *Journal of Economic Theory* 203, 105471.
- Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *The American Economic Review*, 157–177.
- Feenstra, R. C. and M. B. Reinsdorf (2000). An exact price index for the almost ideal demand system. *Economics Letters* 66(2), 159–162.
- Feenstra, R. C. and M. B. Reinsdorf (2007, October). Should Exact Index Numbers Have Standard Errors? Theory and Application to Asian Growth. In *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, NBER Chapters, pp. 483–513. National Bureau of Economic Research, Inc.
- Fisher, F. M. and K. Shell (1968). *Taste and quality change in the pure theory of the true cost of living index*. Value, Capital and Growth. Edinburgh University Press.
- Gerschenkron, A. (1951). *A dollar index of Soviet machinery output, 1927-28 to 1937*, Volume 197. Rand Corporation.
- Hamilton, B. W. (2001, June). Using engel’s law to estimate cpi bias. *American Economic Review* 91(3), 619–630.
- Hanoch, G. (1975). Production and demand models with direct or indirect implicit additivity. *Econometrica*, 395–419.
- Hausman, J. (1981). Exact consumer’s surplus and deadweight loss. *American Economic Review* 71(4), 662–76.
- Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The economics of new goods*, pp. 207–248. University of Chicago Press.
- Herrendorf, B., R. Rogerson, and A. Valentinyi (2013). Two perspectives on preferences and structural trans-

- formation. *American Economic Review* 103(7), 2752–89.
- Hicks, J. R. (1940). The valuation of the social income. *Economica* 7(26), 105–124.
- Hulten, C. R. (1973). Divisia index numbers. *Econometrica: Journal of the Econometric Society*, 1017–1025.
- Hulten, C. R. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies*, 511–518.
- IMF (2004). *Consumer Price Index Manual: Theory and Practice*. USA: International Monetary Fund.
- Ivancic, L., W. E. Diewert, and K. J. Fox (2011). Scanner data, time aggregation and the construction of price indexes. *Journal of Econometrics* 161(1), 24–35.
- Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the us retail sector. *The Quarterly Journal of Economics* 134(2), 715–783.
- Jaravel, X. and D. Lashkari (2022). Nonparametric measurement of long-run growth in consumer welfare. Cepr discussion paper no. dp16866.
- Jones, C. I. and P. J. Klenow (2016). Beyond gdp? welfare across countries and time. *American Economic Review* 106(9), 2426–57.
- Jorgenson, D. W., M. S. Ho, K. J. Stiroh, et al. (2005). productivity, volume 3: information technology and the american growth resurgence. *MIT Press Books* 3.
- Kongsamut, P., S. Rebelo, and D. Xie (2001). Beyond Balanced Growth. *Review of Economic Studies* 68(4), 869–882.
- Konüs, A. A. (1939). The problem of the true index of the cost of living. *Econometrica: Journal of the Econometric Society*, 10–29.
- Lucas, R. E. (1987). *Models of business cycles*, Volume 26. Basil Blackwell Oxford.
- Martin, R. S. (2020). Revisiting taste change in cost-of-living measurement. Technical report, US Department of Labor, US Bureau of Labor Statistics, Office of Prices and Living Conditions.
- Matsuyama, K. (2019). Engel’s law in the global economy: Demand-induced patterns of structural change, innovation, and trade. *Econometrica* 87(2), 497–528.
- McKenzie, G. W. and I. F. Pearce (1982). Welfare measurement: a synthesis. *The American Economic Review* 72(4), 669–682.
- Muellbauer, J. (1975). The cost of living and taste and quality change. *Journal of Economic Theory* 10(3), 269–283.
- Ngai, L. R. and C. A. Pissarides (2007). Structural change in a multisector model of growth. *American economic review* 97(1), 429–443.
- Nordhaus, W. D. (2008). Baumol’s diseases: A macroeconomic perspective. *The BE Journal of Macroeconomics* 8(1), 1–39.
- Oulton, N. (2008). Chain indices of the cost-of-living and the path-dependence problem: An empirical solution. *Journal of Econometrics* 144(1), 306–324.
- Redding, S. J. and D. E. Weinstein (2020). Measuring aggregate price indices with taste shocks: Theory and evidence for CES preferences. *The Quarterly Journal of Economics* 135(1), 503–560.
- Samuelson, P. A. and S. Swamy (1974). Invariant economic index numbers and canonical duality: survey and synthesis. *The American Economic Review* 64(4), 566–593.
- Syverson, C. (2017). Challenges to mismeasurement explanations for the us productivity slowdown. *Journal of Economic Perspectives* 31(2), 165–86.
- United Nations (2009). *System of National Accounts 2008*. USA: United Nations.

Appendix A Proofs

We do not present the proofs in the same order as the propositions in the paper because the proofs of earlier propositions rely on notation and logic from later propositions.

Proof of Lemma 1. By definition,

$$\begin{aligned} EV^m &= \log \frac{e(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}{e(p_{t_0}, v(p_{t_0}, I_{t_0}; x_{t_1}); x_{t_1})} \\ &= \log \frac{e(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}{e(p_{t_0}, v(p_{t_0}, I_{t_0}; x_{t_1}); x_{t_1})} \frac{e(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}{e(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})} \\ &= \log \frac{e(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}{I_{t_0}} \frac{I_{t_1}}{e(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}. \end{aligned}$$

The first equality follows from definition (1), noting that $I_{t_0} = e(p_{t_0}, v(p_{t_0}, I_{t_0}; x_{t_1}); x_{t_1})$ and $\exp(EV^m) \times I_{t_0} = e(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})$. The third equality uses the fact that $I_{t_1} = e(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})$. To finish, rewrite

$$\log \frac{e(p_{t_0}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}{e(p_{t_1}, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})} = - \int_{t_0}^{t_1} \sum_{i \in N} \frac{\partial \log e(p_t, v(p_{t_1}, I_{t_1}; x_{t_1}); x_{t_1})}{\partial \log p_{it}} \frac{d \log p_{it}}{dt} dt,$$

and use Shephard's lemma to express the price elasticity of the expenditure function in terms of budget shares. If the path of prices between t_0 and t_1 is not differentiable, then construct a new modified path of prices that is differentiable, and apply the integral to this modified path. Since the integral is path independent, it only depends on p_{t_0} and p_{t_1} . Therefore any smooth path that connects p_{t_0} and p_{t_1} gives the same integral. \square

Proof of Proposition 1. If the path of prices is continuously differentiable, we can combine Lemma 1 with the definition of real consumption. \square

Proof of Proposition 2. Differentiate real consumption

$$\Delta \log Y = \int_{t_0}^{t_1} \left(\frac{d \log I_s}{ds} - \sum_{i \in N} b_{is} \frac{d \log p_{is}}{ds} \right) ds$$

twice with respect to t_1

$$\begin{aligned} \frac{d \log Y}{dt_1} &= \frac{d \log I}{dt} - \sum_{i \in N} b_i \frac{d \log p_{it}}{dt} \\ \frac{d^2 \log Y}{dt_1^2} &= \frac{d^2 \log I}{dt^2} - \sum_{i \in N} b_i \frac{d^2 \log p_i}{dt^2} - \sum_{i \in N} \frac{db_i}{dt} \frac{d \log p_i}{dt}, \end{aligned}$$

where all right-hand-side terms are evaluated at t_1 . Therefore,

$$\begin{aligned}\Delta \log Y &\approx \frac{d \log Y}{dt_1} \Delta t + \frac{1}{2} \frac{d^2 \log Y}{dt_1^2} (\Delta t)^2 \\ &= \left[\frac{d \log I}{dt} - \sum_{i \in N} b_i \frac{d \log p_{it}}{dt} \right] \Delta t \\ &\quad + \frac{1}{2} \left[\frac{d^2 \log I}{dt^2} - \sum_{i \in N} b_i \frac{d^2 \log p_{it}}{dt^2} - \sum_{i \in N} \frac{db_{it}}{dt} \frac{d \log p_{it}}{dt} \right] (\Delta t)^2.\end{aligned}$$

Thus,

$$\Delta \log Y \approx \Delta \log I - \sum_{i \in N} b_i \Delta \log p_i - \frac{1}{2} \sum_{i \in N} \Delta b_i \Delta \log p_i, \quad (32)$$

where recall that for any variable z , $\Delta \log z \approx (\partial \log z / \partial t) \Delta t + 1/2 (\partial^2 \log z / \partial t^2) (\Delta t)^2$ and the remainder terms are of order $(\Delta t)^3$. Evaluating (32) at $t_1 = t_0$ yields (12).

For EV^m , using Lemma 1, we can write

$$EV^m = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} b_i(p, v(p_{t_1}, x_{t_1}), x_{t_1}) \frac{d \log p_{is}}{ds} ds$$

Differentiate EV^m with respect to t_1

$$\begin{aligned}\frac{dEV^m}{dt_1} &= \frac{d \log I}{dt} - \sum_{i \in N} b_i(p, v(p_{t_1}, x_{t_1}), x_{t_1}) \frac{d \log p_{it}}{dt} \\ &\quad - \int_{t_0}^{t_1} \sum_{i \in N} \frac{du_t}{dt} \frac{\partial b_i(p, v(p_{t_1}, x_{t_1}), x_{t_1})}{\partial u} \frac{d \log p_{is}}{ds} ds - \int_{t_0}^{t_1} \sum_{i \in N} \frac{dx_t}{dt} \frac{\partial b_i(p, v(p_{t_1}, x_{t_1}), x_{t_1})}{\partial x} \frac{d \log p_{is}}{ds} ds,\end{aligned}$$

where the shares and derivatives are evaluated at t_1 , and then differentiate again and evaluate at $t_1 = t_0$ to get

$$\begin{aligned}
\frac{d^2 EV^m}{dt_1^2} &= \frac{d^2 \log I}{dt^2} - \sum_{i \in N} \sum_{j \in N} \frac{\partial b_i}{\partial \log p_i} \frac{d \log p_{it}}{dt} \frac{d \log p_{jt}}{dt} \\
&\quad - \sum_{i \in N} b_i \frac{d^2 \log p_{it}}{dt^2} - 2 \sum_{i \in N} \frac{du_t}{dt} \frac{\partial b_i}{\partial u} \frac{d \log p_{it}}{dt} - 2 \sum_{i \in N} \frac{dx_t}{dt} \frac{\partial b_i}{\partial x} \frac{d \log p_{it}}{dt} \\
&= \frac{d^2 \log I}{dt^2} - \sum_{i \in N} \left[\sum_{j \in N} \frac{\partial b_i}{\partial \log p_j} \frac{d \log p_{jt}}{dt} + \frac{du_t}{dt} \frac{\partial b_i}{\partial u} + \frac{d \log x_t}{dt} \frac{\partial b_i}{\partial x} \right] \frac{d \log p_{it}}{dt} \\
&\quad - \sum_{i \in N} b_i \frac{d^2 \log p_{it}}{dt^2} - \sum_{i \in N} \frac{du_t}{dt} \frac{\partial b_i}{\partial u} \frac{d \log p_{it}}{dt} - \sum_{i \in N} \frac{dx_t}{dt} \frac{\partial b_i}{\partial x} \frac{d \log p_{it}}{dt} \\
&= \frac{d^2 \log I}{dt^2} - \sum_{i \in N} \frac{db_{it}}{dt} \frac{d \log p_{it}}{dt} - \sum_{i \in N} b_i \frac{d^2 \log p_{it}}{dt^2} \\
&\quad - \sum_{i \in N} \frac{du_t}{dt} \frac{\partial b_i}{\partial u} \frac{d \log p_{it}}{dt} - \sum_{i \in N} \frac{dx_t}{dt} \frac{\partial b_i}{\partial x} \frac{d \log p_{it}}{dt},
\end{aligned}$$

where $\partial b_i / \partial \log p_j$ holds fixed utility values u and taste shifters x (without loss of generality, represented by a scalar). The first three terms are equal to the second-order expansion of $\Delta \log Y$, and the remaining terms are the gap between $\Delta \log Y$ and EV^m . Finally, we use $db_i = \sum_j d \log p_j \frac{\partial b_i}{\partial \log p_j} + du \frac{\partial b_i}{\partial u} + dx \frac{\partial b_i}{\partial x}$ to obtain (13). \square

Proof of Proposition 3. Substitute (14) in place of Δb in (32) to get the expression for $\Delta \log Y$. For the gap between real consumption and EV^m , note that Proposition 2 implies that

$$EV^m - \Delta \log Y \approx -\frac{1}{2} \sum_i \left[\Delta b_i - \sum_j \frac{\partial b_i}{\partial \log p_j} \Delta \log p_j \right] \Delta \log p_i$$

Using (14) in place of Δb above and the fact that $\partial b_i / \partial \log p_i = (1 - \theta_0) b_i (1 - b_i)$ for $i = j$ and $\partial b_i / \partial \log p_j = \theta_0 b_i b_j$ for $i \neq j$, yields the following

$$\Delta \log EV^m - \Delta \log Y \approx -\frac{1}{2} \sum_{i \in N} \left[(\varepsilon_i - 1) b_i \left(d \log I - \sum_{j \in N} b_j \Delta \log p_j \right) + b_i \Delta \log x_i \right] \Delta \log p_i,$$

which can be rearranged to give the desired expression. \square

Proof of Proposition 6. For this proof, we use notation introduced in Section 4. By Shep-

hard's lemma, for each $i \in N$,

$$d \log p_i = -d \log A_i + \sum_{j \in N+F} \Omega_{ij} d \log p_j.$$

For $i \in F$, letting nominal GDP be the numeraire, we have

$$d \log p_i + d \log A_i + d \log L_i = d \log \Lambda_i,$$

where Λ is the vector of factor income shares, $d \log A_i$ represents a factor-quality shock, and $d \log p_i$ for a factor is the change in the factor price per efficiency unit. Inverting this linear system yields

$$d \log p = -\Psi d \log A + \Psi^F (d \log \Lambda - d \log L),$$

where Ψ^F is the $(1 + N + F) \times F$ matrix consisting of the right F columns of Ψ . For $i \in F$, this equation is

$$d \log p_i = -d \log A_i + d \log \Lambda_i - d \log L_i.$$

The change in the wage per factor unit (not efficiency units) is $d \log p_i + d \log A_i$.

We can write (simplifying notation by suppressing dependence on t in the integral)

$$\begin{aligned} \Delta \log Y &= - \int_{t_0}^{t_1} b' d \log p \\ &= - \int_{t_0}^{t_1} b' \left[-\Psi d \log A - \Psi^F d \log L + \Psi^F d \log \Lambda \right] \\ &= \int_{t_0}^{t_1} \lambda' d \log A + \int_{t_0}^{t_1} \Lambda' d \log L - \int_{t_0}^{t_1} \Lambda d \log \Lambda \\ &= \int_{t_0}^{t_1} \lambda' d \log A + \int_{t_0}^{t_1} \Lambda' d \log L \end{aligned}$$

where the second line substitutes in our solution for $d \log p$, and the third line uses the fact that $\lambda' = b' \Psi$, $\Lambda' = b' \Psi^F$, and the fourth line uses $\Lambda' d \log \Lambda = 0$ because the factor shares always sum to one: $\sum_{f \in F} \Lambda_f = 1$. \square

Proof of Proposition 5. Recall that the macro equivalent variation at final preferences is defined by $EV^M = \phi$, where

$$V(A_{t_0}, e^\phi L_{t_0}; x_{t_1}) = V(A_{t_1}, L_{t_1}; x_{t_1})$$

Denote by $p(A, L, x)$ goods prices under technologies A , factor quantities L , and prefer-

ences x . Without loss of generality, we fix income at I . We have $p_{t_1} \equiv p(A_{t_1}, L_{t_1}, x_{t_1})$ and

$$u_{t_1} \equiv v(p_{t_1}, I; x_{t_1}) = V(A_{t_1}, L_{t_1}; x_{t_1}).$$

Define a hypothetical Hicksian economy with fictional households that have stable homothetic preferences defined by the expenditure function $e^h(p, u) = e(p, u_{t_1}; x_{t_1})u$ where e is the expenditure function of the consumer in the primitive economy. Budget shares of this fictional consumer are $b_i^h(p) \equiv \frac{\partial \log e^h(p, u)}{\partial \log p_i} = \frac{\partial \log e(p, u_{t_1}; x_{t_1})}{\partial \log p_i}$. Given any technology vector, in this hypothetical Hicksian economy we denote the Leontief inverse matrix by Ψ^h and sales shares by λ^h . Given technologies A_t and factor quantities L_t , we denote prices in this hypothetical economy by p_t^h . Changes in prices in this hypothetical economy satisfy

$$d \log p^h = -\Psi^h d \log A + \Psi^{hF} d \log \Lambda^h, \quad (33)$$

where Ψ^h is the fictitious Leontief inverse. Note that $p(A_{t_1}, L_{t_1}, x_{t_1}) = p^h(A_{t_1}, L_{t_1})$ and $p(A_{t_0}, e^\phi L_{t_0}, x_{t_1}) = p^h(A_{t_0}, e^\phi L_{t_0})$, where we used the fact that $V(A_{t_0}, e^\phi L_{t_0}; x_{t_1}) = u_{t_1}$. We will use the property that, with constant returns to scale, homothetic preferences, and constant income I ,

$$p^h(A, aL) = \frac{1}{a} p^h(A, L)$$

for every $a > 0$. Using the previous results,

$$\begin{aligned} V(A_{t_0}, e^\phi L_{t_0}; x_{t_1}) &= v(p(A_{t_0}, e^\phi L_{t_0}, x_{t_1}), I; x_{t_1}) \\ &= v(p^h(A_{t_0}, e^\phi L_{t_0}), I; x_{t_1}) \\ &= v(e^{-\phi} p^h(A_{t_0}, L_{t_0}), I; x_{t_1}) \\ &= v(p^h(A_{t_0}, L_{t_0}), e^\phi I; x_{t_1}), \end{aligned}$$

where the last equality used the fact that the value function is homogeneous of degree zero in prices and income. We thus have

$$v(p^h(A_{t_0}, L_{t_0}), e^\phi I; x_{t_1}) = v(p^h(A_{t_1}, L_{t_1}), I; x_{t_1}),$$

which can be re-expressed using the expenditure function as

$$EV^M = \log \frac{e(p^h(A_{t_1}, L_{t_1}), u_{t_1}; x_{t_1})}{e(p^h(A_{t_0}, L_{t_0}), u_{t_1}; x_{t_1})}.$$

This observation is a key step in the proof. Macro welfare changes can be re-expressed as

micro welfare changes given changes in equilibrium prices in a fictional Hicksian economy with preferences represented by $e^h(p, u)$. As in the proof of Lemma 1, rewrite EV^M as (suppressing dependence on t in the integral)

$$EV^M = - \int_{t_0}^{t_1} \sum_{i \in N+F} \frac{\partial \log e(p, u_{t_1})}{\partial \log p_i} d \log p_i^h = - \int_{t_0}^{t_1} \sum_{i \in N+F} b_i^h d \log p_i^h.$$

Following the same steps as in the proof of Proposition 6 (for the hypothetical economy), we obtain

$$EV^M = \int_{t_0}^{t_1} \sum_{i \in N} \lambda_i^h d \log A_i + \int_{t_0}^{t_1} \sum_{f \in F} \lambda_f^h d \log L_f.$$

□

Proof of Proposition 4. By the proof of Proposition 5, $EV^m = EV^M$ if, and only if, $p^h(A_t, L_t) = p(A_t, L_t, x_t)$ for all t , where $p^h(A_t, L_t)$ is defined in Proposition 5. This condition is immediate if preferences are homothetic and stable. Consider now the case in which preferences are non-homothetic and/or unstable but factor income shares, Λ , are constant (using notation introduced in Section 4 and in the proof of Proposition 6). Then by Proposition 8, changes in prices in response to changes in A , L , and x are given by the following differential equation:

$$d \log p = -\Psi d \log A - \Psi^F d \log L.$$

Furthermore, note that changes in Ψ are determined by changes in Ω since $\Psi = (I - \Omega)^{-1}$. Since every $i \in N$ has constant returns to scale, changes in Ω_{ij} depend only on changes in relative prices for every $i \in N$. This means that changes in Ω only depend on changes in relative prices, therefore changes in Ψ depend only on changes in relative prices. Since x and utility u do not appear in any of these expressions, this means that prices and incomes $p(A, L, x)$ and $I(A, L, x)$, relative to the numeraire, do not depend on x and u . Thus, $p^h(A_t, L_t) = p(A_t, L_t, x_t)$. □

Proof of Proposition 7. The proof follows similar steps as in the proof of Proposition 3. Write changes in real GDP (abstracting from changes in factor endowments) as,

$$\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} \lambda_{is} \frac{d \log A_i}{ds} ds,$$

differentiate twice with respect to t_1 and evaluate the derivative at $t_1 = t_0$. Write EV^M as

$$EV^M = \int_{t_0}^{t_1} \sum_{i \in N} \lambda_i(A_s, L_s, u_{t_1}, x_{t_1}) \frac{d \log A_i}{ds} ds,$$

differentiate twice with respect to t_1 and evaluate the derivative at $t_1 = t_0$. This yields the desired expression. \square

Proof of Proposition 8. Following the steps in the proof of Proposition 6, for each $i \in N$ we have

$$d \log p_i = -d \log A_i + \sum_{j \in N+F} \Omega_{ij} d \log p_j,$$

and for each $i \in F$ (recalling that we assume that factor endowments are unchanged and normalizing nominal GDP to be the numeraire) we have

$$d \log p_i = -d \log A_i + d \log \Lambda_i,$$

where $d \log A_i$ represents a factor-quality shock when i is a factor. Combining these yields the desired expression for changes in prices

$$d \log p = -\Psi d \log A + \Psi^F d \log \Lambda.$$

To get changes in sales shares, note that

$$\begin{aligned} \lambda &= b' \Psi \\ d\lambda &= d(b' \Psi) \\ &= b' \Psi d\Omega \Psi + db' \Psi \end{aligned}$$

$$\Omega_{ij} d \log \Omega_{ij} = (1 - \theta_i) \Omega_{ij} (d \log p_j - \sum_k \Omega_{ik} d \log p_k)$$

$$d\Omega_{ij} = (1 - \theta_i) \text{Cov}_{\Omega_{(i,:)}}(d \log p, Id_{(:,j)})$$

$$\begin{aligned} \sum_j d\Omega_{ij} \Psi_{jk} &= (1 - \theta_i) \text{Cov}_{\Omega_{(i,:)}}(d \log p, Id_{(:,j)}) \Psi_{jk} \\ &= (1 - \theta_i) \sum_j \text{Cov}_{\Omega_{(i,:)}}(d \log p, \Psi_{jk} Id_{(:,j)}) \end{aligned}$$

$$[d\Omega \Psi]_{ik} = (1 - \theta_i) \text{Cov}_{\Omega_{(i,:)}}(d \log p, \Psi_{(:,k)})$$

Meanwhile

$$\begin{aligned}
d \log b_i &= (1 - \theta_0) \left(d \log p_i - \sum_i b_i d \log p_i \right) + (\varepsilon_i - 1) d \log Y + d \log x_i \\
&= (1 - \theta_0) \text{Cov}_{\Omega_{(0,:)}} \left(d \log p, Id_{(:,i)} \right) + \text{Cov}_{\Omega_{(0,:)}} \left(\varepsilon, Id_{(:,i)} \right) d \log Y + \text{Cov}_{\Omega_{(0,:)}} \left(d \log x, Id_{(:,i)} \right) \\
\sum_i db_i \Psi_{ik} &= \text{Cov}_{\Omega_{(0,:)}} \left((1 - \theta_0) d \log p + \varepsilon d \log Y + d \log x, \Psi_{(:,k)} \right)
\end{aligned}$$

Hence, $d\lambda' = \lambda' d\Omega\Psi + db'\Psi$ can be written as

$$d\lambda_k = \sum_i \lambda_i (1 - \theta_i) \text{Cov}_{\Omega_{(i,:)}} (d \log p, \Psi_{(:,k)}) + \text{Cov}_{\Omega_{(0,:)}} \left(\varepsilon, \Psi_{(:,k)} \right) d \log Y + \text{Cov}_{\Omega_{(0,:)}} \left(d \log x, \Psi_{(:,k)} \right).$$

□

Proof of Proposition 9. Normalize nominal GDP to one. Applying Proposition 8 to a one-factor model yields

$$d \log p = -\Psi d \log A,$$

so that relative prices do not respond to changes in demand or income. To solve for $\Delta \log Y$, use Proposition 7 in combination with the expression for $d \log p$ and $d\lambda$ in Proposition 8 in the case of one factor. To solve for EV^M , by Proposition 4, $EV^M = EV^m$. Solve for $EV^m - \Delta \log Y$ by plugging the expression for $d \log p$ into Proposition 2 and noting that $b' = \Omega_{(0,:)}$. □

Proof of Proposition 10. By Lemma 1, welfare changes measured as the equivalent variation at t_1 preferences (which is equal to compensating variation at t_1 tastes since preferences are homothetic) is given by

$$EV^m = \log \frac{I_{t_1}}{I_{t_0}} - \log \frac{P_{t_1}}{P_{t_0}}, \quad (34)$$

where

$$\frac{P_{t_1}}{P_{t_0}} = \left(\frac{\sum_i x_{it_1} p_{it_0}^{1-\theta_0}}{\sum_j x_{jt_1} p_{jt_1}^{1-\theta_0}} \right)^{\frac{1}{\theta_0-1}} = \left(\frac{\sum_i x_{it_1} p_{it_1}^{1-\theta_0} \left(\frac{p_{it_0}}{p_{it_1}} \right)^{1-\theta_0}}{\sum_j x_{jt_1} p_{jt_1}^{1-\theta_0}} \right)^{\frac{1}{\theta_0-1}}. \quad (35)$$

We re-express (35) as

$$\begin{aligned} \left(\frac{P_{t_1}}{P_{t_0}}\right)^{\theta_0-1} &= \frac{\sum_{i \in \mathcal{C}} x_{it_1} p_{it_1}^{1-\theta_0} \left(\frac{p_{it_0}}{p_{it_1}}\right)^{1-\theta_0} + \sum_{i \in \mathcal{N}} x_{it_1} p_{it_1}^{1-\theta_0} \left(\frac{p_{it_0}}{p_{it_1}}\right)^{1-\theta_0} + \sum_{i \in \mathcal{X}} x_{it_1} p_{it_0}^{1-\theta_0}}{\sum_j x_{jt_1} p_{jt_1}^{1-\theta_0}} \quad (36) \\ &= b_{t_1}^c \sum_{i \in \mathcal{C}} b_{it_1}^c \left(\frac{p_{it_0}}{p_{it_1}}\right)^{1-\theta_0} + (1 - b_{t_1}^c) \sum_{i \in \mathcal{N}} b_{it_1}^n \left(\frac{p_{it_0}}{p_{it_1}}\right)^{1-\theta_0} + b_{t_0}^{x_{t_1}} \left(\frac{P_{t_1}}{P_{t_0}}\right)^{\theta_0-1} \end{aligned}$$

where $b_{t_1}^c \equiv \sum_{i \in \mathcal{C}} b_{it_1}$ denotes the expenditure share on continuing goods at t_1 ,

$$b_{it_1}^c \equiv \frac{x_{it_1} p_{it_1}^{1-\theta_0}}{\sum_{i \in \mathcal{C}} x_{it_1} p_{it_1}^{1-\theta_0}}, \quad b_{it_1}^n \equiv \frac{x_{it_1} p_{it_1}^{1-\theta_0}}{\sum_{i \in \mathcal{N}} x_{it_1} p_{it_1}^{1-\theta_0}},$$

and

$$b_{t_0}^{x_{t_1}} \equiv \frac{\sum_{i \in \mathcal{X}} x_{it_1} p_{it_0}^{1-\theta_0}}{\sum_j x_{jt_1} p_{jt_0}^{1-\theta_0}}$$

is the (unobserved) share of exiting goods at t_0 prices under t_1 preferences. Defining a price index for continuing goods under t_1 preferences,

$$\frac{P_{t_1}^c}{P_{t_0}^c} = \left(\sum_{i \in \mathcal{C}} b_{it_1}^c \left(\frac{p_{it_1}}{p_{it_0}}\right)^{\theta_0-1} \right)^{\frac{1}{\theta_0-1}}$$

and a price index for new goods under t_1 preferences,

$$\frac{P_{t_1}^n}{P_{t_0}^n} = \left(\sum_{i \in \mathcal{N}} b_{it_1}^n \left(\frac{p_{it_1}}{p_{it_0}}\right)^{\theta_0-1} \right)^{\frac{1}{\theta_0-1}},$$

we can express the change in the welfare-relevant price index as

$$\frac{P_{t_1}}{P_{t_0}} = \left(\frac{b_{t_1}^c \left(\frac{P_{t_1}^c}{P_{t_0}^c}\right)^{\theta_0-1} + (1 - b_{t_1}^c) \left(\frac{P_{t_1}^n}{P_{t_0}^n}\right)^{\theta_0-1}}{1 - b_{t_0}^{x_{t_1}}} \right)^{\frac{1}{\theta_0-1}}. \quad (37)$$

□