

Growth Accounting with Demand Instability and Income Effects

David R. Baqaee
UCLA

Ariel Burstein*
UCLA

April 7, 2021

Abstract

We study how welfare responds to changes in budget and production possibility sets when preferences are unstable or non-homothetic. We characterize the gap between welfare and standard measures of real consumption in partial and general equilibrium. We show that whenever changes in expenditure shares due to non-homotheticity or instability covary with changes in prices (in partial equilibrium) or changes in technology (in general equilibrium), measures of real consumption are biased because they do not properly account for consumer substitution. We apply our results to long-run and short-run phenomena. In the long-run, we show that structural transformation, if caused by non-homotheticities, has much larger implications for welfare than what is implied by standard measures of Baumol's cost disease. In the short-run, we show that when firms' demand shocks are correlated with their supply shocks, there is a gap between measured and welfare-relevant industry-level output that does not disappear in the aggregate. Finally, we show that correlated supply and demand shifters make real GDP an unreliable metric for measuring production, and illustrate this using the Covid-19 crisis.

*We thank Sihwan Yang for superb research assistance. We thank Andy Atkeson, Natalie Bau, 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.

1 Introduction

How does a change in the economic environment affect welfare? For example, how does the welfare of a consumer change when the budget constraint changes, or how does the welfare of a nation change when the production possibility frontier of the economy changes? At first blush, answering this question seems very difficult, perhaps requiring detailed information about the nonlinear functions describing preferences and technologies.

Under the very strong assumption of homotheticity and preference stability, standard theory offers a simple non-parametric procedure for recovering the answer. Consider a change in prices and income that occurred over some time horizon t_0 to t_1 . Indexing individual goods by i , the change in welfare is approximately equal to

$$\Delta \text{Welfare} \approx \Delta \log I - \sum_{t=t_0}^{t_1} \sum_i b_i(t) (\log p_i(t+1) - \log p_i(t)), \quad (1)$$

where $\Delta \log I$ is the log change in nominal income, and $b_i(t)$ and $p_i(t)$ are the budget share and price of good i at time t . In words, the change in nominal income deflated by the expenditure-share weighted change in prices approximates the change in welfare. The fact that the expenditure shares are updated at every period t between t_0 and t_1 reduces substitution bias, and eliminates it in the continuous-time limit.

Equation (1) is a *chain-linked* index, and such indices are used to measure most types of real economic activity, ranging from aggregates like real GDP, total factor productivity (TFP), private consumption and investment, to less aggregated objects like industry-level measures of inflation and production. The fact that these indices approximate changes in welfare and production under homotheticity and stability justifies their recommended use in the United Nations' System of National Accounts.¹

While homotheticity and preference stability are highly convenient assumptions, they have counterfactual implications: homotheticity requires that income effects are uniform, that is, the income elasticity of demand equals one for every good; stability requires that consumers only change their spending decisions in response to changes in incomes and relative prices. In this paper, we provide sufficient statistics that adjust equation (1) when these assumptions are relaxed.² Our baseline welfare measure is the equivalent variation

¹See OECD et al. (2004) and references therein, in particular Chapters 15 and 17, for a comprehensive overview and discussion of price and quantity indices and their relation to welfare.

²As we discuss in detail in Section 2, preference instability is driven by any factor that changes preference rankings over bundles of goods at fixed prices and income, e.g. age, health, advertising, fads. In the literature, preference instability and non-homotheticities are typically studied independently. We analyze

at fixed final preferences, which answers the question: “holding fixed prices and preferences, how much income must consumers be given to make them indifferent between their choice sets at t_0 and t_1 ?”

We study this problem in both partial equilibrium, where choice sets are defined in terms of budget sets (prices and income are exogenous), and in general equilibrium, where choice sets are defined in terms of production possibility frontiers (prices and income are endogenous).³ We provide exact and approximate characterizations of the adjustment to equation (1), and demonstrate that the failure of chain-linked indices, in both partial and general equilibrium, is caused by the fact that they *undermeasure* the substitution caused by either non-homotheticity or preference instability. That is, with non-homotheticities or preference instability, chained indices suffer from exactly the type of substitution bias that they were designed to eliminate. We show that this bias is larger if changes in sales shares caused by non-homotheticity or taste shocks are correlated with changes in prices (in partial equilibrium) or changes in technology (in general equilibrium). If these correlations are zero, then no adjustment is required and (1) holds.

Our partial equilibrium welfare measure answers a microeconomic question for an infinitesimal agent who cannot alter market-level prices through her choices. Our general equilibrium welfare measure answers a macroeconomic question for a collection of agents whose collective decisions alter market-level prices. When preferences are stable, macroeconomic changes in welfare are equal to microeconomic changes in welfare. However, we show that these two measures are not equal when household preferences change over time. Intuitively, when household preferences change over time, a comparison of budgets through time is not the same as a comparison of production possibility sets because, using the consumer’s terminal preferences, the consumer may make a choice on the initial budget constraint that is not on the original production possibility frontier.

Our results for welfare and the gap between welfare and real consumption are expressed in terms of measurable sufficient statistics. Surprisingly, in both partial and general equilibrium, we show that computing the change in welfare caused by changes in prices (in partial equilibrium) or technology (in general equilibrium) does not require direct knowledge of the taste shocks or the shape of non-homotheticities. Instead, what we must know are the expenditure shares and the elasticities of substitution at the final allocation. For the micro problem, these are the household expenditure shares and the elasticities of substitution in consumption. For the macro problem, these are the input-

them jointly in this paper because both generate the same type of biases in measures of real consumption. Our results are relevant when either of these forces is active.

³For the macro problem, we consider neoclassical economies with representative agents.

output table and the elasticities of substitution in both production and consumption.

Our general equilibrium results show how the details of the production structure, like input-output linkages, complementarities in production, and decreasing returns to scale, interact with non-homotheticities and preference shocks to magnify the gap between welfare and GDP. Furthermore, we show that there is no reason to expect the discrepancies between welfare and GDP that we emphasize to get aggregated away. This means that the correction necessary to bring GDP in line with welfare will, in general, depend greatly on the details of production possibility set, and can get larger the more we disaggregate. In this sense, our results are related to the literature studying the macroeconomic implications of production networks and disaggregation (e.g. Gabaix, 2011; Acemoglu et al., 2012; Baqaee and Farhi, 2019b).

We illustrate the relevance of our results for understanding short-run and long-run phenomena by means of three applications. In our first application, we analyze the importance of non-homotheticity or preference instability for measures of long-run productivity growth. Since Baumol (1967), an enduring stylized fact about economic growth has been the observation that industries with slow productivity growth tend to become larger as a share of the economy over time. This phenomenon, known as Baumol's cost disease, means that as the industrial mix of the economy changes in favor of low-productivity-growth industries, aggregate growth is increasingly determined by productivity growth in the slow-growth industries. To be specific, from 1948 to 2014, aggregate TFP in the US grew by 60%. If the US economy had kept its original 1948 industrial structure, then aggregate TFP would have grown by 78% instead. We show that if structural transformation is caused solely by non-homotheticity and demand instability, then welfare-relevant TFP grew by only 47%. This is because measured aggregate TFP does not fully account for substitution caused by changes in demand, and hence, the increase in the welfare-relevant measure of aggregate TFP is much lower than what is measured.

In our second application, we consider a firm-level specification of our model. We show that when firms' demand shocks are correlated with their supply shocks, there is a gap between welfare-relevant and measured changes in industry-level output and prices. We show that these biases, which can be sizable even at annual frequency, do not disappear as we aggregate up to the level of real GDP even if firms and industries are infinitesimal. At annual frequency, the gap between welfare and real GDP due only to firm-level supply and demand shocks could be as high as 1%, and this gap gets larger at lower frequencies if firm-level supply and demand shocks are persistent, becoming unboundedly large in the limit when the shocks are random walks. If we start with industry-level (rather than firm-level) data, we are ruling out the existence of these biases by construc-

tion.

In our final application, we show that if changes in household spending patterns are in part driven by demand shifters, then real GDP can become an unreliable metric for measuring changes in production. What should be irrelevant details, like the order of supply and demand shocks, can cause real GDP to be different between the initial and final periods even if initial and final prices and quantities are the same. To illustrate these results, we consider the large changes in household spending patterns towards low-contact goods and services over the first few months of 2020 due to the Covid-19 pandemic. Since these changes in spending patterns were not entirely driven by market prices, we show that real GDP is path dependent. For example, if the path of supply and demand shocks during the recovery does not look exactly like the initial collapse in reverse, measured real GDP can be as much as 6% higher or lower than what it was before the crisis, even if every price and quantity returns to its pre-Covid value.

Of course, there are other reasons, besides instability and non-homotheticity, why (1) can fail to accurately measure welfare. Many of the well-known reasons why the approximation in (1) fails can be thought of as being due to missing prices and quantities. For example, it is well-known that (1) fails to properly account for the creation and destruction of goods if we cannot 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); equation (1) also fails to properly account for changes in the quality of goods (see Syverson, 2017); finally, (1) fails to properly account for changes in non-market components of welfare, like changes in leisure and mortality (see Jones and Klenow, 2016), or changes in the user cost of durables. In all of 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. In this paper, we abstract from these 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. That is, the corrections we derive are different to the ones that are equivalent to adjustments in prices.⁴

⁴Our approach to calculate ex-post welfare changes requires well-measured prices and elasticities of substitution in the final period. For ex-post welfare measurement, when information on prices is missing or mismeasured, if preferences are non-homothetic an alternative approach is to infer 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 e.g. Hamilton, 2001 and, more recently, Atkin et al., 2020). In addition to ex-post measurement with non-homotheticities, in this paper we study the implications of instability of preferences (that generate shifts in expenditures correlated with price changes) and we also consider counterfactuals.

Relatedly, preference instability and mismeasured prices (i.e. unobserved quality change) are sometimes viewed as alternative means to the same end. This is because they can both be used to justify why demand curves shift over time, even holding prices and incomes fixed. That is, both can rationalize changes in behavior that are not triggered by changes in observed prices or incomes. However, while they have similar implications for changes in prices and quantities, they have very different implications for welfare. When there are unobserved changes in quality, the gap between welfare and real consumption is caused by a difference between measured and welfare-relevant *prices*. We show that in the case of non-homotheticities and taste shocks, the gap between welfare and real consumption is caused by a difference between measured and welfare-relevant *expenditure shares*.

Other related literature. This paper contributes to the literatures on growth and productivity accounting, multi-sectoral and disaggregated macroeconomics, as well as the literature on structural transformation. We discuss the way our paper complements and relates to these literatures in turn.

A key assumption in growth accounting is the existence of a stable and homothetic final aggregator. As shown by, for example Hulten (1973) among others, chain-linked indices are meaningful if, and only if, a homothetic and stable final aggregator exists. Therefore, this assumption is ubiquitous in growth accounting, and also appears in almost all papers that study aggregate outcomes using disaggregated input-output models.⁵ We show that when preferences are non-homothetic or unstable, then the “ideal” price deflator is shock-dependent. This allows us to provide a generalization of Domar (1961) and Hulten (1978) that measures changes in welfare in situations when preferences are unstable or non-homothetic. Using this, we can construct exact and approximate characterizations of how welfare responds to shocks in general equilibrium.

Our approach contrasts with the one in Redding and Weinstein (2020). They show that variations in sales are difficult to explain via shifts in supply curves alone, and shifts in demand curves are an important source of variation in the data. They interpret changes in demand curves as being due to changes in tastes, but unlike us, they treat changes in tastes as being equivalent to changes in price. Operationally, this makes the taste shocks behave like quality shocks. They estimate changes in taste/quality necessary to explain variations in product-level data. However, this only determines changes in the *relative* size of demand shocks across goods, and it does not pin down changes in the overall level of these shocks. Redding and Weinstein (2020) pin down the overall level of the shocks by assuming they are mean zero. They then derive CES price indices for changes

⁵See, for example, the review paper by Carvalho and Tahbaz-Salehi (2018) and the references therein.

in utility in the presence of such shocks. Our approach is different in that we do not compare utils before and after the taste shocks. Instead we compute changes in equivalent variation keeping preferences over goods constant for the variation, as advocated by Fisher and Shell (1968) and Samuelson and Swamy (1974). This approach does not require any assumptions about the overall level of the taste shocks in terms of utils.⁶

Our paper is also related to the literature on structural transformation and Baumol's 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 non-homotheticities, or income effects, whereby households spend more of their income on some goods as they become richer (e.g. Kongsamut et al., 2001; Boppart, 2014; Comin et al., 2015; Alder et al., 2019).

Our results suggest that settling this question has important implications for welfare. From a welfare perspective, structural transformation driven by relative price changes should be treated differently to structural transformation driven by non-homotheticity or demand instability. In particular, measures of real production or consumption must be adjusted for substitution bias in the latter case, but no adjustment is necessary in the former case.

The structure of the paper is as follows. In Section 2, we set up the microeconomic problem and provide exact and approximate characterizations of the difference between welfare and measured real consumption changes. In Section 3, we set up the macroeconomic general equilibrium model and provide exact and approximate characterizations of the difference between welfare and measured 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 some extensions in Section 6 and conclude in Section 7. All proofs and additional details are presented in an appendix which will be included in the next

⁶Some papers in index number theory have studied the relationship between conventional index numbers and welfare in the presence of preference instability and non-homotheticities. For example, Caves et al. (1982) show that when preferences are homothetic, translog, but unstable, measured Tornqvist-based indices correspond to a geometric average of welfare changes under initial and final preferences. A similar result holds in our context, as we show in Appendix B, but in the body of the paper our focus is different. We characterize welfare (in partial and general equilibrium) at either initial or final preferences and using either EV or CV, rather than averaging these different measures.

draft.

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 these measures of welfare with measures of real consumption. We provide exact and approximate results. We model the equilibrium determination of prices in Section 3.

2.1 Definition of Welfare and Real Consumption

In this subsection we define welfare and real consumption. Measuring changes in welfare using equivalent variation is standard when preferences are stable, even when they are non-homothetic. However, measuring welfare changes in the presence of unstable preferences is less common and therefore we discuss this issue in some detail.

Consider a set of preference relations, $\{\succeq_x\}$, over bundles of goods. These preferences are indexed by x , which represents anything that affects preference rankings over bundles of goods. For example, x could be calendar time, age, exposure to advertising, geographic location, or state of nature. For every x , we represent the preference relation \succeq_x by a utility function $u(c; x)$, where c is an $N \times 1$ vector and N is the number of goods in the consumption bundle. Since the consumer makes no choices over x , we do not need to specify how $u(c; x)$ varies with x . Moreover, preferences over x , if they exist, are not revealed by choices.⁷

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

Definition 1 (Homotheticity). Preferences over goods c are *homothetic* if, for every positive scalar $a > 0$ and every feasible c and x , we can write

$$u(ac; x) = au(c; x).$$

Definition 2 (Stability). Preferences over goods c are *stable* if there exists a time-invariant function $\Phi(\cdot)$ such that the utility function can be written as $u(c; x) = U(\Phi(c); x)$ for every feasible c and x .

⁷In Section 6, we discuss situations in which x is endogenously chosen and valued by the consumer, such as leisure, but its price and quantity are not being measured. We also discuss situations in which x is endogenously chosen by firms, such as advertising.

If preferences are stable, x can change over time (e.g. households get higher or lower utils from all goods) but, since x is separable from c , these changes do not impact preferences over bundles of goods c . If preferences are not stable, we say that they are unstable.

Given preferences encapsulated in u , the indirect utility function of the consumer, 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 income. Since utility is only defined up to monotone transformations, changes in utility do not have meaningful units. Therefore, we measure welfare in money terms by means of the expenditure function. For any value of x , the expenditure function is given by

$$e(p, u; x) = \min_c \{p \cdot c : u(c; x) = u\}.$$

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} . This change in the budget set is accompanied by changes in x from x_{t_0} to x_{t_1} . Our baseline measure of welfare is defined as follows.

Definition 3 (Micro Welfare). The change in welfare measured using the *micro equivalent variation* with *final preferences*:

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

EV^m measures the change in income, under initial prices p_{t_0} , that a consumer with final preferences, $u(c; x_{t_1})$, would need to be indifferent between the initial budget set defined by (p_{t_0}, I_{t_0}) and the new budget set defined by (p_{t_1}, I_{t_1}) .⁸ The superscript m represents the fact that this is the *micro* equivalent variation, since we take prices as given.

Figure 1 illustrates EV^m in a simple two good economy. Point A depicts the consumer's choice at some initial equilibrium. Point B shows the consumer's new choice in response to a shock to both the consumer's budget set and preferences (the indifference curve associated with utility $v(p_{t_1}, I_{t_1}; x_{t_1})$). When preferences change, the new and old indifference curve intersect. Point C shows the choice the consumer would have made

⁸In principle, we could also measure changes in welfare using compensating (instead of equivalent) variation, or by using initial (rather than final) preferences (see e.g. Balk, 1989 for a discussion of the various ways one can define welfare changes). Combining EV with final preferences (CV with initial preferences) is natural since this requires preserving the shape of the indifference curve at the final (initial) allocation. In the body of the paper, for conciseness, we focus on EV using final preferences since equivalent variation is more commonly used and final preferences are more relevant than initial preferences, but we characterize these other welfare measures in Appendix B. See also Remark 1.

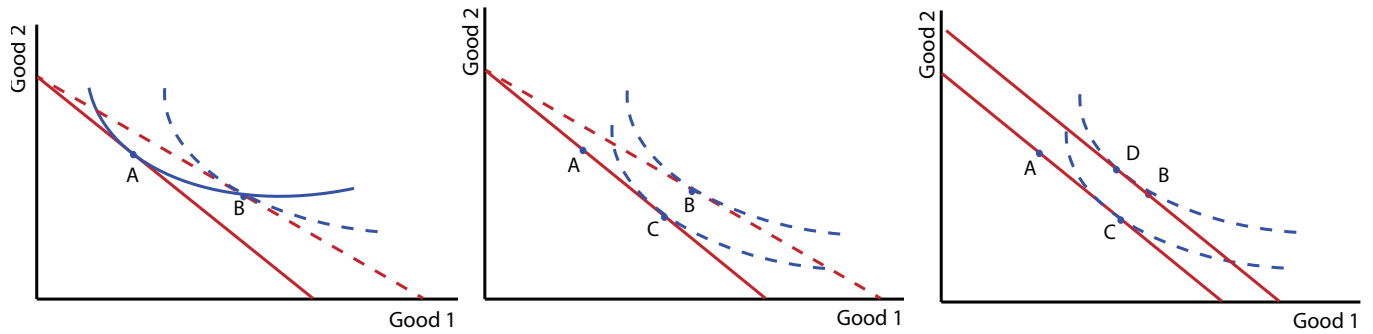


Figure 1: Graphical depiction of the micro equivalent variation using final preferences. The panel on the left is the initial and final equilibrium. Indifference curves cross due to taste shocks. The panel in the middle is how we compare the original and final budget sets taking into account non-homotheticities and preference instability. The panel on the right is how we calculate the equivalent variation

in the old budget set under the new preferences (the indifference curve associated with utility $v(p_{t_0}, I_{t_0}; x_{t_1})$). Point D is the point at which the old budget line is tangent to the indifference curve containing point B (with utility $v(p_{t_1}, I_{t_1}; x_{t_1})$). The micro equivalent variation, EV^m , is the log distance between the initial budget constraint (at points A and C) and the budget constraint at point D.⁹

Discussion of our welfare criterion. As pointed out by Fisher and Shell (1968), the welfare criterion in Definition 3 is different to an alternative one that would attempt to measure the change in income that a consumer would need to be as well off in t_0 as in t_1 . Whereas the former refers to a choice that the consumer can in principle make, the latter criterion relies on an arbitrary intertemporal comparison of utils which do not correspond to any choice.¹⁰

To be concrete, suppose that x represents the age of the consumer. Clearly, we cannot meaningfully compare the amount of utils an individual derives from playing with toys during childhood to the amount of utils that same individual derives from drinking wine during adulthood. Since consumers never make choices about how old they are, their

⁹In other words, we evaluate the welfare effect of changes in constraints (from one budget set to another), not the welfare effects of changes in allocations (from point A to point B).

¹⁰By an intertemporal comparisons of utils, we do not refer to dynamic decisions made at a point in time. To see this, imagine a dynamic decision maker with an intertemporal utility function. Suppose that after every period, we multiply the entire intertemporal utility function of the consumer by a large positive number. Doing this has no observable implications, since at every point in time, the consumer faces the exact same trade-offs as if we had not multiplied their utility. However, it dramatically alters the amount of utils received in one period versus another. Note that assumptions on time discounting (which discipline saving decisions) are about future versus present payoffs evaluated at a given point in time, so they do not pin down the level utils in one time period versus the level utils in another time period.

preferences across consumption goods consumed at different ages are not revealed by their choices. In the words of Heraclitus: “No man ever steps in the same river twice, for it’s not the same river and he’s not the same man.” A comparison of utils during childhood to utils during adulthood is as meaningless as a comparison of utils between two different individuals. On the other hand, if we fix the consumer’s age x , we can meaningfully compare the consumer’s preferences about the bundles of goods they consumed at different points in their life or different consumption streams that they may face in the future.

This approach, of holding x constant, is different to the one taken when x represents some form of quality change (usually taken to be isomorphic to an unobserved reduction in prices). Intuitively, quality adjustments are more applicable to situations where the consumer can conceivably make choices between the good at differing levels of quality. For example, if a box of chocolates undergoes quality change so that each box now contains twice as many chocolates, the consumer can conceivably make choices between the old and new boxes that reveal how much they value the quality change. Taste changes, on the other hand, do not involve meaningful choices from the consumer’s perspective — if a consumer decides that she prefers dark chocolate to white chocolate, it does not make sense to ask how she would trade off consuming white chocolate in the past, when she preferred white chocolate, to consuming dark chocolate in the present, when she prefers dark chocolate. Instead, holding fixed her preferences, we can ask how she trades off white chocolate against dark chocolate. Therefore, the welfare implications of changes in unmeasured quality and changes in tastes are very different.

Of course, this is not to say that quality changes do not happen in reality. Rather, that we will abstract from quality change in our analysis. If there is quality change that can be represented as an unobservable price reduction, then all of our results will apply to the quality-adjusted “correctly-measured” prices instead of the market prices.

Real Consumption. Having defined changes in welfare, we now define changes in real consumption. The change in real consumption corresponds to what national income accountants and statistical agencies do when given data on the evolution of prices p and consumption bundles c . We assume that this data is *perfect* — completely accurate, comprehensive, adjusted for any necessary quality changes, and available in continuous time. This is because the biases associated with imperfections in the data, like the lack of quality adjustment, missing prices, or infrequent measurement, are different to the biases we study.

Definition 4 (Real consumption). For some continuously differentiable path of prices that

unfold as a function of some scalar t (interpreted as time), the change in *real consumption* from t_0 to t_1 is defined to be

$$\Delta Y = \int_{t_0}^{t_1} \sum_{i \in N} p_i(t) \frac{dc_i}{dt} dt. \quad (2)$$

That is, changes in real consumption are cumulated changes in consumption goods measured at constant instantaneous prices. Equation (2) is called a *Divisia* quantity index (in honor of François Divisia). In practice, since perfect data is not available in continuous time, statistical agencies approximate this integral via a (Riemann) sum using chained indices (e.g. Fisher or Tornqvist). We abstract from the imperfections of these approximations in this paper.

In log terms, (2) can be rewritten as

$$\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} b_i(t) \frac{d \log c_i}{dt} dt = \int_{t_0}^{t_1} \sum_{i \in N} b_i d \log c_i, \quad (3)$$

where $b_i(t)$ is the budget share of good i given prices, income, and preferences at time t . The last equation on the right-hand side simplifies notation by suppressing dependence on t in the integral.¹¹ Using the budget constraint, we can express changes in real consumption in terms of changes in income deflated by price changes,

$$\Delta \log Y = \Delta \log I - \int_{t_0}^{t_1} \sum_{i \in N} b_i d \log p_i = \Delta \log I - \Delta \log P^Y. \quad (4)$$

In other words, changes in real consumption are equal to changes in income minus changes in the consumption price deflator. Notice that changes in real consumption (or 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.

2.2 Relating Welfare and Consumption

We consider how real consumption and welfare change in response to changes in the budget set and the indifference curves of the consumer. We first consider global results and then local results. The results are stated in terms of changes in prices and income,

¹¹For any variable z , we denote by dz its change over infinitesimal time intervals, so that $\Delta z = \int_{t_0}^{t_1} dz$.

which we endogenize in Sections 3 and 4.

Global results. Received wisdom from consumer theory is that non-homothetic preferences do not have ideal price indices associated with them. That is, when preferences are non-homothetic or unstable, there does not exist a function of prices that converts changes in income into changes in welfare.

When preferences are unstable or non-homothetic, the following lemma shows that income must be deflated by a shock-dependent price index. Changes in this price index are equal to budget-share weighted price changes, as in expression (4) for real consumption. However, whereas the price deflator for real consumption is based on observed budget shares (given prices, income, and preferences over time), the price deflator for welfare is based on hypothetical budget shares (at fixed utility level and fixed preferences).

To state this, define the budget share of good i (given prices, preferences and a level of utility) by

$$b_i(p, u; x) \equiv \frac{p_i c_i(p, u; x)}{e(p, u; x)} = \frac{\partial \log e(p, u; x)}{\partial \log p_i}, \quad (5)$$

where the second equality, Shephard's lemma, establishes a connection between budget shares and elasticities of the expenditure function. Note that when preferences are homothetic, then the expenditure function is separable in utility $e(p, u; x) = e(p; x) u$ and, hence, budget shares do not depend on u .

The following lemma characterizes changes in microeconomic welfare.

Lemma 1 (Micro Welfare). *Given any change in prices, income, and preferences, micro welfare changes are given by*

$$EV^m = \Delta \log I - \int_{t_0}^{t_1} \sum_i b_i^{ev} d \log p_i = \Delta \log I - \Delta \log P^{EV}, \quad (6)$$

where $b_i^{ev} \equiv b_i(p, v(p_{t_1}, I_{t_1}, x_{t_1}), x_{t_1})$ denotes budget shares at prices p , fixed final preferences, and fixed final utility.

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

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

and recognizing that the second term can be written as the integral in (6).

Compared to real consumption (4), which weights price changes by observed budget shares, welfare weights price changes by hypothetical budget shares evaluated at current

prices but for fixed final preferences and final utility. Intuitively, budget shares respond to relative price changes, changes in preferences, and changes in utility. However, for welfare, we must only take into account changes in budget shares that occurred due to the changes in relative prices (fixing utility and preference parameters).

Equivalently, we can reinterpret the hypothetical budget shares b^{ev} as corresponding to those of a fictional consumer with homothetic and stable preferences with expenditure function $e^*(p, u) = e(p, v^*; x_{t_1}) \frac{u}{v^*}$, where $v^* = v(p_{t_1}, I_{t_1}, x_{t_1})$. Figure 1 helps explain the intuition: to measure EV , we find the level of income at point D at which the *final* indifference curve becomes tangent to the *initial* budget line – this requires preserving the shape of the indifference curve at the final allocation (rather than the initial one).¹²

Lemma 1 implies that we can measure changes in welfare given changes in prices *without* needing to know income elasticities or the nature of demand shocks. What we must know, instead, are the terminal budget shares and constant-utility elasticities of substitution at the terminal equilibrium. That is, all we need to know are the elasticities of substitution of $e(p, u_{t_1}; x_{t_1})$.^{13,14}

Remark 1 (Compensating Variation under Initial Preferences). Our baseline measure of welfare changes is equivalent variation under final preferences. An alternative would be to use compensating variation under initial preferences. Every result in the paper can be translated into compensating variation under initial preferences simply by reversing the flow of time. In particular, whereas Lemma 1 preserves the shape of the indifference curve at the final allocation, the compensating variation counterpart to Lemma 1 preserves the shape of the indifference curve at the initial allocation. Hence, calculating compensating variation requires knowledge of budget shares and elasticities of substitution at the initial allocation, whereas equivalent variation requires knowledge of budget shares and elasticities of substitution at the final allocation. See Appendix B for more details.¹⁵

¹²As indicated by equation (6), we can measure EV as the distance in the y-intercept between the initial budget line tangent to D and the final budget line tangent to B (plus the change in income, which equals zero in this example). Specifically, this vertical distance is equal to EV^m in levels: $\exp(EV^m) = \Delta I - \int_{t_0}^{t_1} \sum_i \frac{\partial e(p, v_{t_1}, x_{t_1})}{\partial p_i} dp_i = \Delta I - \int_{t_0}^{t_1} \sum_i c_i(p, v_{t_1}, x_{t_1}) dp_i$.

¹³In particular, if we know these elasticities of substitution (globally), then we can integrate the expression in Lemma 1. If all of these substitution elasticities are equal to one globally, then Lemma 1 implies that the welfare-relevant price deflator is the Paasche index. Auer et al. (2021) apply Lemma 1 to measure the heterogeneous welfare effects of changes in foreign prices in the presence of demand non-homotheticities.

¹⁴These elasticities of substitution can be backed out from the price elasticity of Hicksian demand. If highly disaggregated data is available, where the budget shares of individual goods are approximately zero, cross-price elasticities of Hicksian demand can be estimated without knowledge of income elasticities since income effects are negligible in this case. If income effects are non-negligible, then knowledge of income elasticities will be necessary to recover Hicksian elasticities (which are necessary for welfare calculations) from Marshallian elasticities (which are estimable).

¹⁵In Appendix B we show that, up to a second-order approximation (but not globally), changes in real

Combining Lemma 1 with equation (4) results in the following proposition, which compares changes in welfare and real consumption.

Proposition 1 (Consumption vs. Welfare). *Given any change in prices, income, and preferences, the difference between welfare changes and real consumption is*

$$EV^m - \Delta \log Y = \int_{t_0}^{t_1} \sum_i (b_i - b_i^{ev}) d \log p_i = (t_1 - t_0) \mathbb{E}_t \text{Cov}(b - b^{ev}, d \log p),$$

where the covariance is calculated across goods at a point in time, and the expectation is calculated across time between t_0 and t_1 .

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 changes in utility u over time. Hence, whenever preferences are homothetic and stable, $b_i^{ev}(t) = b_i(t)$ for every path of shocks and every t . In other words, we have the following corollary.

Corollary 1 (Homothetic and Stable Preferences). *Welfare changes equal real consumption, $EV^m = \Delta \log Y$, if, and only if, preferences are homothetic and stable.*

Since changes in welfare depend only on initial and final values of primitives, Corollary also 1 implies that real consumption is path-independent whenever preferences are homothetic and stable.

When preferences are non-homothetic or unstable, observed budget shares not only reflect price changes but also non-price changes (that is, changes in x and changes in u). This generates discrepancies between observed and hypothetical budget shares, and hence between real consumption and welfare.

For a given path of price and income changes, welfare exceeds real consumption if, on average, $\text{Cov}(b - b^{ev}, d \log p) < 0$. That is, if changes in expenditure shares due to changes in u or x favor goods whose prices are falling more rapidly. Welfare changes are smaller than changes in real consumption if this pattern is reversed. If deviations between expenditure shares and relative price movements are orthogonal, then real consumption and welfare are equal. For example, with Cobb-Douglas preferences where expenditure shares are subject to exogenous shocks, real consumption and welfare are equal if, and only if, shocks to prices and shocks to consumer preferences are orthogonal.

consumption equal a simple average of equivalent variation under final preferences and compensating variation under initial preferences.

The global expressions for real consumption and welfare in Lemma 1 and Proposition 1 require knowledge of the full path of observed and hypothetical budget shares. In what follows we use a second-order approximation around initial choices to characterize changes in real consumption and welfare. These expressions are helpful for gaining intuition and for later calculations.

Local results. We consider local approximations of the objects of interest as the time period goes to zero, $t_1 - t_0 = \Delta t \rightarrow 0$. Throughout the rest of the paper, a second-order approximation means that the remainder term is of order Δt^3 . We focus on second-order approximations to capture the interaction between price changes and expenditure-switching, which is the source of the gaps between real consumption and welfare changes. To clarify the intuition about the gap between real consumption and welfare, we begin by stating the results in terms of Hicksian budget shares, and then we re-express them in terms of Marshallian (observable) budget shares.

We start by characterizing the change in real consumption.

Lemma 2 (Approximate Consumption). *Up to a second order approximation, the change in real consumption is*

$$\Delta \log Y \approx \Delta \log I - b' \Delta \log p - \frac{1}{2} \sum_{i \in N} \left[\Delta \log p' \frac{\partial b_i}{\partial \log p} + \Delta \log x' \frac{\partial b_i}{\partial \log x} + \Delta \log v \frac{\partial b_i}{\partial \log u} \right] \Delta \log p.$$

This lemma, which is standard, shows that a second-order approximation accounts for the fact that budget shares change over time. The first term in the square brackets reflects changes in budget shares due to changes in relative prices and the next two terms correspond to changes in budget share due to non-price factors, preferences (under unstable preferences) and utility (under non-homothetic preferences).¹⁶

As discussed above, welfare is measured using changes in budget shares at fixed final preferences and utility. Thus, when comparing our welfare measures to real consumption we must add in changes in budget shares due to non-price factors, as indicated in the following proposition:

Proposition 2 (Approximate Welfare vs. Consumption). *To a second-order approximation,*

¹⁶The terms $\Delta \log x$ and $\Delta \log u$ need only be first-order approximations since they are multiplied by $\Delta \log p$ (and we only need to keep terms that are of order Δt^2). However, only for the first term $-b' \Delta \log p$, the primitive shock in prices must be approximated up to the second order, that is, $\Delta \log p \approx (\partial \log p / \partial t) \Delta t + 1/2 (\partial^2 \log p / \partial t^2) \Delta t^2$.

the change in welfare is given by

$$EV \approx \Delta \log Y - \frac{1}{2} \sum_{i \in N} \left[\Delta \log x \frac{\partial b_i}{\partial \log x} + \Delta \log u \frac{\partial b_i}{\partial \log u} \right] \Delta \log p. \quad (7)$$

To a first order, changes in welfare equal real consumption. To understand the (second-order) gap between welfare and real consumption changes, consider first the case of homothetic but unstable preferences. Whereas changes in real consumption only take into consideration changes in budget shares in response to changes in utility parameters as the shock unfolds over time, changes in welfare must account for these changes from the start. In other words, real consumption does not sufficiently account for substitution caused by preference instability. For example, the additional reduction in welfare (at new preferences) from a price increase in a good i with increasing demand ($d \log x \frac{\partial b_i}{\partial \log x} d \log p_i > 0$) is not fully reflected in real consumption, implying $EV^m < \Delta \log Y$.¹⁷

Similar reasoning applies in the case of stable but non-homothetic preferences, since changes in budget shares due to non-homotheticities should be incorporated in welfare immediately but are reflected in real consumption only gradually. For example, a reduction in the price of a good for which income effects are relatively weak ($d \log v \frac{\partial b_i}{\partial \log v} d \log p_i > 0$) implies a smaller increase in welfare than in real consumption ($EV^m < \Delta \log Y$).

Lemma 2 and Proposition are both expressed in terms of Hicksian elasticities. We now re-express these results in terms of Marshallian elasticities. For clarity of exposition, we assume that household preferences are determined by a non-homothetic CES aggregator as in Comin et al. (2015) or Fally (2020).¹⁸ In this case, a loglinearization of household budget shares yields

$$d \log b_i = (1 - \theta_0) (d \log p_i - \mathbb{E}_b[d \log p]) + (\varepsilon_i - 1) (d \log I - \mathbb{E}_b[d \log p]) + d \log x_i, \quad (8)$$

where $\mathbb{E}_b(\cdot)$ is the expectation using budget shares as probability weights. All terms in this expression are, at least in principle, observable. The elasticity ε_i is the income elasticity of good i , and θ_0 is the constant-utility elasticity of substitution across goods. The parameter θ_0 can be disciplined by the fact that $(\theta_0 - \varepsilon_i)b_j$ is the price elasticity of (Marshallian) demand for good i relative to j 's price. These elasticities are evaluated at

¹⁷A non-zero correlation between prices and demand shifters may emerge endogenously if firms have non-constant returns to scale or if firms invest in advertisement in response to productivity shocks. We consider the first possibility in Example 4 in Section 4 and discuss the second in Section 6.

¹⁸Our results can be generalized to arbitrary non-CES functional forms, but since the intuition for the more general case is very similar to the CES case, we leave the more general non-parametric results in Appendix C.

the initial point around which we approximate. The term $d \log x_i$ is a demand shifter, a residual that captures changes in observable expenditures not attributable to changes in income or prices.¹⁹ Note that when ε_i is equal to 1 for every i , final demand is homothetic, and when x_i is constant for all i , final demand is stable.

Proposition 3 (Approximate Micro using Marshallian Demand). *Consider some perturbation in final demand $\Delta \log x$, prices $\Delta \log p$, and income $\Delta \log I$. Then, to a second-order approximation, the change in real consumption is*

$$\begin{aligned} \Delta \log Y \approx & \Delta \log I - \mathbb{E}_b [\Delta \log p] - \frac{1}{2}(1 - \theta_0) \text{Var}_b (\Delta \log p) \\ & - \frac{1}{2} \text{Cov}_b (\Delta \log x, \Delta \log p) - \frac{1}{2} (\Delta \log I - \mathbb{E}_b [\Delta \log p]) \text{Cov}_b (\varepsilon, \Delta \log p), \end{aligned} \quad (9)$$

and the change in welfare is

$$EV \approx \Delta \log Y - \frac{1}{2} \text{Cov}_b (\Delta \log x, \Delta \log p) - \frac{1}{2} (\Delta \log I - \mathbb{E}_b [\Delta \log p]) \text{Cov}_b (\varepsilon, \Delta \log p), \quad (10)$$

where $\text{Cov}_b(\cdot)$ is the covariance using the initial budget shares as the probability weights.

We begin by considering the change in real consumption in (9). To a first order, the change in real consumption is just the change in income deflated by prices: $\Delta \log I - \mathbb{E}_b[\Delta \log p]$. The remaining terms capture nonlinearities associated with expenditure-switching. Since these are second-order, they are multiplied by 1/2. We discuss these terms one-by-one. If goods are substitutes, $\theta_0 > 1$, then variance in relative prices boosts expenditure shares of cheaper goods and this increases measured real consumption. The second line captures the changes due to changes in demand. Intuitively, if the composition of demand shifts in favor of goods that happen to become relatively cheap, either due to non-homotheticity $\text{Cov}_b(\varepsilon, \Delta \log p) (\Delta \log I - \mathbb{E}_b[\Delta \log p]) < 0$ or demand shocks $\text{Cov}_b(\Delta \log x, \Delta \log p) < 0$, then real consumption increases.

Now consider changes in welfare in (10). As expected, the first-order terms are identical. The remaining terms capture the nonlinear response of welfare to price shocks. Note that if preferences are stable and homothetic, then welfare changes coincide with changes in real consumption. However, if preferences are unstable or non-homothetic, real consumption strays from welfare whenever price changes covary with non-price changes in

¹⁹Since b_i are expenditure shares that always add up to one, it must necessarily be the case that $\mathbb{E}_b[d \log x] = 0$, and $\mathbb{E}_b[\varepsilon] = 1$. See Appendix G for a derivation of this log-linearization of Marshallian demand.

demand. This happens because real consumption “undercounts” expenditure-switching due to the changes in demand.

3 Macroeconomic Changes in Welfare and Consumption

In the previous section we showed how changes in budget sets (prices and income) affect welfare when preferences are unstable and non-homothetic. For these problems, the frontier of the consumer’s choice set is linear. This means that relative prices do not respond to the choices of the consumer. At the level of a whole society however, choice sets are not linear. The production possibility set associated with an economy may have a nonlinear frontier. In this case, relative prices respond endogenously to choices made by consumers. In this section, we extend our analysis to allow for nonlinear choice sets — that is, for nonlinear production possibility frontiers (PPFs). The analysis in this section collapses to the one in Section 2 when the production possibility frontier of the economy is the same as the budget constraint (as happens in very simple general equilibrium models).

We first update our definitions of welfare, now at the macroeconomic level, and we introduce some basic structure and notation. We then present expressions for real GDP and welfare at the macroeconomic level, first globally and then locally in terms of endogenous sufficient statistics. In the next section, Section 4, we solve for these endogenous objects in terms of observable primitives.

3.1 Definition of Welfare and Real GDP

Consider a perfectly competitive neoclassical production economy with a representative agent.²⁰ The production possibilities of the economy are defined by a production possibility set, $G(c; A) \leq 0$, where c is the vector of goods the economy delivers and A is a set of parameters that controls endowments and production technologies. For simplicity, we refer to A as *technology* with the understanding that it includes factor endowments.

By the first welfare theorem, the allocation is at the efficient frontier, satisfying $G(c; A) = 0$. This means that the equilibrium price of good i relative to j is equal to the slope of the PPF $\frac{p_i}{p_j} = \frac{\partial G / \partial c_i}{\partial G / \partial c_j}$. This is in contrast to our micro results, which took prices as given and did not require that the first welfare theorem holds.

²⁰In order for the macro welfare problem to be meaningful, we require both that household preferences be aggregable (so that there is a structural representative agent), and that the equilibrium be Pareto efficient. We discuss non-aggregable preferences in Section 6. If the equilibrium is not Pareto-efficient, then we cannot value changes in choice sets via revealed preferences since society is not collectively optimizing. When the decentralized equilibrium is inefficient or preferences are non-aggregable, we can still rely on the micro welfare change defined in Section 2, which requires neither assumption.

Define the *macro indirect utility* function as the maximum amount of utility the competitive equilibrium can deliver

$$v(A; x) = \max\{u(c; x) : G(c; A) = 0\}.$$

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 technology as the primitive and lets society pick feasible points in the production set, with prices determined in general equilibrium.

Definition 5 (Macro Welfare). The change in welfare associated with a change in technology sets ΔA is measured using the *macro equivalent variation with final preferences*,

$$EV^M = \log \frac{e(p_{t_0}, v(A_{t_1}; x_{t_1}); x_{t_1})}{e(p_{t_0}, v(A_{t_0}; x_{t_1}); x_{t_1})}.$$

The distinction between micro and macro welfare only becomes relevant when preferences are unstable. If preferences are stable, then the macro welfare change associated with ΔA coincides with the micro welfare changes associated with the $(\Delta p, \Delta I)$ that ΔA causes. In other words, when preferences are stable, suppressing dependence on x , we have

$$EV^M = \log \frac{e(p_{t_0}, v(A_{t_1}))}{e(p_{t_0}, v(A_{t_0}))} = \log \frac{e(p_{t_0}, v(p_{t_1}, I_{t_1}))}{e(p_{t_0}, v(p_{t_0}, I_{t_0}))} = EV^m.$$

This follows from the fact that $v(A_{t_1}) = v(p_{t_1}, I_{t_1})$ and $v(A_{t_0}) = v(p_{t_0}, I_{t_0})$. However, micro and macro welfare do not coincide when preferences are unstable. To see this, consider the two-good example depicted in Figure 2.

Point A in the left panel depicts the allocation under initial technologies and initial preferences (the indifference curve associated with utility $v(A_{t_0}; x_{t_0})$). The budget line going through point A is based on initial prices, given by the slope of the PPF at that point. Point B in the center panel shows allocations under final technologies and final preferences (the indifference curve associated with utility $v(A_{t_1}; x_{t_1})$). The budget line going through point B is based on final prices given by the slope of the PPF at that point. Point C shows the choice of consumers with final preferences facing a budget constraint defined by initial prices and income. This point C is the micro welfare change, exactly as in Figure 1. Note however that point C is not feasible at the initial technologies (it is outside the PPF). Instead, point D depicts the best feasible allocation under initial technologies and final preferences, with resulting utility $v(A_{t_0}; x_{t_1})$. Note that prices associated to point D are different to those associated to points B and C. Finally, EV^M is given by the distance

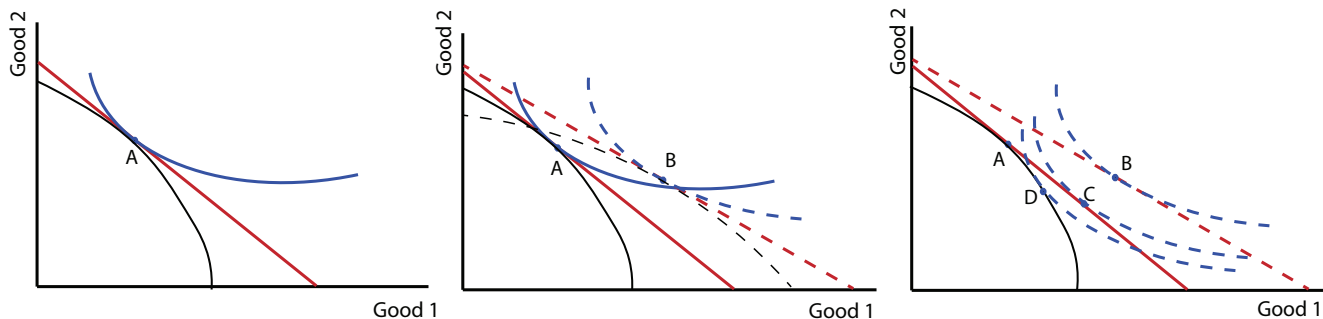


Figure 2: Graphical depiction of the macro and micro equivalent variation using final preferences. The panel on the left is the initial equilibrium. The panel in the middle is the final equilibrium. The panel on the right shows the macro and micro choices of the consumer, with initial preferences, in the initial equilibrium. Point C is what they would pick on the original budget set, whereas point D is what they would pick on the original production possibility frontier.

between the budget line (under initial prices) tangent to the indifference curve associated with point D relative to the one associated with point B.

As in Section 2, to study this problem we index the path of technology shocks and preference shocks, $\log A(t)$ and $\log x(t)$, by some scalar t that represents time. The definition of ΔY is the same as before: $\Delta Y = \int_{t_0}^{t_1} \sum_{i \in N} p_i dc_i$. In the general equilibrium model and (its applications), we refer to ΔY as real GDP in this economy since it coincides with the chain-weighted change in real GDP.²¹

3.2 Relating Welfare and Real GDP

We now characterize changes in real GDP and welfare, first globally and then locally. The results in this subsection are the general equilibrium counterparts to those in Section 2. They are “reduced-form” in the sense that they are not expressed in terms of primitives. In Section 4, we explicitly solve for these sufficient statistics in terms of observable primitives.

To analyze the macroeconomic problem, we spell out the production structure of the

²¹We abstract from international trade, government spending, and investment, all of which drive a gap between nominal GDP and nominal consumption. The most rigorous way to deal with investment is to cast the model in intertemporal terms and index goods by period of time in which they are consumed, along the lines of Basu et al. (2012). Theoretically, this presents no issues and our results apply to such economies unchanged. In practice, this changes the way we would map our model to the data, where nominal GDP would now be defined in net-present-value terms and prices would have to be multiplied by the (potentially stochastic) discount factor. A simpler way to justify our static application of these formulas is to either assume that consumption and investment bundles of goods are the same (so the consumer values investment goods in the same way as they value consumption goods) or to treat investment as a static intermediate input (assuming full depreciation every period).

economy in more detail. Suppose that 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 G_i is a neoclassical production function, m_{ij} are intermediate inputs used by i and produced by j , and l_{if} denotes factor inputs used by i . Factors of production $f = 1, \dots, F$ are inelastically supplied. The exogenous scalar A_i is a Hicks-neutral productivity shifter. Without loss of generality, we assume that G_i has constant returns to scale because decreasing returns to scale is isomorphic to adding producer-specific factors. The technology vector A is of length $(N + F)$: the first N elements are productivities for each $i \in N$, and the last F elements are factor supplies for each $f \in F$.

Define the sales shares relative to GDP of each good $i \in N$ to be

$$\lambda_i = \frac{p_i y_i}{I}.$$

These are often referred to as *Domar* weights. Note that referring λ_i as a “share” is an abuse of language since $\sum_{i \in N} \lambda_i > 1$ whenever there are intermediate inputs.

Global Results. The following results show that changes in real GDP and welfare can both be represented as sales-weighted averages of technology changes. Real GDP uses actual sales shares over time, while welfare uses sales shares in an artificial economy in which budget shares only respond to price changes.

Lemma 3 (Real GDP). *Given a change in technologies and preferences, the change in real GDP is*

$$\Delta \log Y = \int_{t_0}^{t_1} \sum_{i \in N} \lambda_i d \log A_i, \quad (11)$$

where λ are sales shares which are functions of A and x .

Lemma 3 shows that changes in real GDP are equal to sales-weighted changes in technology. This is a slight generalization of Hulten (1978) to environments with unstable and non-homothetic final demand.

Next, we show that a Hulten-style result also exists for changes in welfare. Define $\lambda^{ev}(A, x)$ to be sales shares in a fictional economy with the same production possibility set but where consumers have stable homothetic preferences represented by the expenditure function $e(p, u) = e(p, v^*, x_{t_1})u$ where $v^* = v(A_{t_1}, x_{t_1})$, similar to Section 2.

Proposition 4 (Macro Welfare). *Changes in macro welfare are*

$$EV^M = \int_{t_0}^{t_1} \sum_{i \in N} \lambda_i^{ev} d \log A_i, \quad (12)$$

if $e(p, v(A, x), x)$ is either homothetic or stable. Furthermore, the difference between changes in welfare and real GDP are

$$EV^M - \Delta \log Y = (t_1 - t_0) \mathbb{E}_t \text{Cov}(\lambda^{ev} - \lambda, d \log A). \quad (13)$$

Note that, in contrast to the micro welfare results in Lemma 1, Proposition 4 requires that preferences are either stable or homothetic. In Appendix F, we discuss the complications that arise when preferences are both non-homothetic and unstable. Whenever the PPF is linear, as it would be in general equilibrium models with a single factor of production, then this requirement on preferences is unnecessary (see Proposition 5 below).

According to Proposition 4, growth accounting for welfare should be based on hypothetical sales shares evaluated at current technology but for fixed final preferences and final utility. The derivation of (12), when preferences are stable but non-homothetic, follows from the fact that $EV^M = EV^m$ so that Lemma 1 can be used in combination with the observation that the budget-share weighted changes in prices at every point is equal to the sales-share weighted changes in productivities. When preferences are unstable but homothetic, the result follows from recognizing that holding fixed preferences, changes in real GDP will measure changes in welfare, and so Hulten's theorem can be applied in a counterfactual economy where preferences are held constant at their final level. The detailed derivations are in the appendix, here, we discuss some salient implications of this proposition.

The first implication is that for welfare questions, the only information we need about preferences are elasticities of substitution at the final allocation, since the fictional consumer in Proposition 4 has stable preferences with income elasticities all equal to one.²²

Second, Proposition 4 shows that real GDP is equal to the change in welfare if, and only if, preferences are homothetic and stable (in which case $\lambda = \lambda^{ev}$ for every level of A and x). That is, Corollary 1, introduced for our micro welfare notion, holds in general equilibrium.

Third, as stated in the following corollary, movements on the surface of a PPF driven by changes in preferences have no effect on macroeconomic welfare or real GDP.

²²Following the observation made in Remark 1, for compensating variation at initial preferences, we need to know elasticities of substitution at the initial allocation instead of the final one.

Corollary 2 (Demand Shocks Only). *In response to changes in preferences, x , that keep the technology vector unchanged, $A(t) = A(t_0)$ for $t \in [t_0, t_1]$,*

$$\Delta \log Y = EV^M = 0.$$

However, micro welfare changes, EV^m , may be nonzero.

As the economy moves along the surface of a PPF, sales shares and prices change, but these changes have no effect on real GDP. This is because demand-driven changes in output raise some quantities and reduce others, and these effects exactly cancel out. Furthermore, since the production possibility set is not changing, macro welfare also does not change. Whereas macro welfare changes EV^M are zero, micro welfare changes EV^m are typically non-zero when prices change. These results are not contradictory: the micro welfare metric assumes that consumers can choose any bundle in their budget set at given prices (hence welfare changes as prices change), whereas the macro welfare metric takes into account changes in prices as the economy chooses different points along the production possibility frontier.

While real GDP and macroeconomic welfare changes are the same so long as we stay on the surface of a given PPF, the two are not equal when the PPF shifts. This is because real GDP is based on a path of sales shares λ that take into consideration technology shocks as well as changes in preferences and non-homotheticities in final demand. However, changes in welfare are based on a path of sales shares λ^{ev} that only take into consideration technology shocks. According to equation (13), if productivity rises for goods in which sales shares fall due to non-technological factors, then $EV^M < \Delta \log Y$.

As mentioned earlier, in general, macro and micro welfare changes are not the same. However, under some conditions, the two coincide.

Proposition 5 (Macro vs. Micro Welfare). *Consider an economy that is subject to technology shocks. Then, macro and micro welfare changes are equal ($EV^m = EV^M$) if (i) preferences are stable, or (ii) preferences are unstable but factor income shares are constant (as in a one factor economy).*

To understand this, consider a change in technologies giving rise to some change in prices. Under condition (i), with stable preferences, households facing a linear budget at new prices (micro welfare) choose the same bundle that maximizes welfare given the fully non-linear production possibility frontier (macro welfare). In contrast, if preferences are not stable, micro and macro welfare changes may not be equal, as we already discussed in the context of demand shocks only in Corollary 2. This is because the choice of consumers

at initial prices and final preferences may not be feasible at the economy-wide level, and hence may differ from the choice that maximizes welfare at the initial technology set and final preferences. This is exactly the difference between points C and D illustrated by Figure 2. However, if the PPF is linear, which is equivalent to condition (ii), micro and macro welfare changes must be equal even under unstable preferences.

Propositions 4 and 5 require as inputs the path of observed and hypothetical sales shares. In the following section, we use a second-order approximation to characterize changes in real GDP and welfare without knowledge of the full path of sales shares.

Local Results. We characterize, up to a second order approximation (as $t_1 - t_0 = \Delta t \rightarrow 0$), the response of real GDP and welfare to technology and preference shocks, now taking into account the endogenous evolution of sales shares.

Lemma 4 (Approximate Real GDP). *Up to a second order approximation, the change in real GDP is²³*

$$\Delta \log Y \approx \lambda' \Delta \log A + \frac{1}{2} \sum_{i \in N} \left[\Delta \log x' \frac{\partial \lambda_i}{\partial \log x} + \Delta \log A' \frac{\partial \lambda_i}{\partial \log A} \right] \Delta \log A_i. \quad (14)$$

Equation (14) resembles the one in Lemma 2, but it is based on sales shares and technology shocks rather than budget shares and price changes. The first term in (14) corresponds to the Hulten-Domar formula. The terms in square brackets reflect 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 due to substitution effects. This type of effect, known as Baumol's cost disease, is an important driver of the slow-down in aggregate productivity growth.

The following proposition compares the change in welfare with the change in real GDP.

²³In equation (14), the ij th element of $\frac{\partial \lambda}{\partial \log x}$ and $\frac{\partial \lambda}{\partial \log A}$ are $\frac{\partial \lambda_j}{\partial \log x_i}$ and $\frac{\partial \lambda_j}{\partial \log A_i}$, respectively.

Proposition 6 (Approximate Macro Welfare vs. GDP). *Up to a second order approximation,*²⁴

$$EV^M \approx \Delta \log Y + \frac{1}{2} \sum_{i \in N} \left[\Delta \log x' \frac{\partial \lambda_i}{\partial \log x} + \Delta \log A' \frac{\partial \log v}{\partial \log A} \frac{\partial \lambda_i}{\partial \log v} \right] \Delta \log A_i, \quad (15)$$

*if either preferences are homothetic or stable.*²⁵

The intuition underlying the gap between macro welfare and real GDP in Proposition 6 is similar to that in Proposition 2 for our micro results. 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. However, welfare measures treat changes in shares due to technology shocks differently than changes in shares due to demand shocks or non-homotheticities. In both cases, real GDP “undercorrects” for changes in shares caused by non-homotheticities or changes in preferences. In particular, welfare is lower than real GDP if technology growth is lower in goods where budget shares rise due to preference changes or non-homotheticities. This observation has important implications for the welfare implications of Baumol’s cost disease, since structural transformation due to demand-side forces should not be treated in the same way as those induced by changes in relative prices. In Section 5, we show that this may imply that welfare has grown considerably less than real GDP over time.

4 Structural Macro Results and Analytic Examples

The results in the previous section are reduced-form in the sense that they take changes in prices and sales shares as given and are written using compensated (Hicksian) demand. In this section, we solve for changes in these endogenous objects in terms of observable sufficient statistics. We first consider economies with linear PPFs and then economies with nonlinear PPFs. We provide some analytical examples to provide more intuition. Although our results are local, one can use them to conduct global counterfactuals, as we do in Section 5, along the lines of Baqaee and Farhi (2019a).

²⁴In equation (15), the ij th element of $\frac{\partial \lambda}{\partial \log x}$ is $\frac{\partial \lambda_j}{\partial \log x_i}$. $\frac{\partial \log v}{\partial \log A}$ is a $(1 + N + F) \times 1$ vector, $\frac{\partial b}{\partial \log v}$ is a $1 \times (1 + N + F)$ vector and $\frac{\partial \log \Lambda}{\partial \log A}$ is a $(1 + N + F)^2$ matrix whose ij th element is $\frac{\partial \Lambda_j}{\partial \log A_i}$. Finally, $d \log p$ is a $(1 + N + F) \times 1$ vector whose first element is the marginal cost of producing the final good c , the next N elements are changes in goods prices for each $i \in N$, and the last F elements are the change in factor prices for each $f \in F$.

²⁵When preferences are non-homothetic but stable $\Delta \log A' \frac{\partial v}{\partial \log A} \frac{\partial \lambda}{\partial v} \Delta \log A = \Delta \log A' \frac{\partial v}{\partial \log A} \frac{\partial b}{\partial v} (-\Delta \log p)$.

We now spell out a macroeconomic model and solve for changes in prices and shares in general equilibrium. For clarity, we restrict attention to nested-CES economies. The general case is in Appendix C, and the intuition is very similar.

Nested-CES economies. Nested-CES economies can be written in many different equivalent ways, since they may have arbitrary patterns of nests. We adopt the following representation. Household preferences are represented by a non-homothetic CES aggregator, which imply that budget shares vary according to (8). We assume that each good $i \in N$ is produced with the production function

$$\frac{y_i}{\bar{y}_i} = \frac{A_i}{\bar{A}_i} \left(\sum_{j \in N+F} \bar{\Omega}_{ij} \left(\frac{m_{ij}}{\bar{m}_{ij}} \right)^{\frac{\theta_i-1}{\theta_i}} \right)^{\frac{\theta_i}{\theta_i-1}},$$

where m_{ij} are intermediate inputs (including factors) used by i and produced by j . Any nested-CES production network can be represented in this way if we treat each CES aggregator as a separate producer (see Baqaee and Farhi, 2019b).

Factor shares. Define the factor income shares of this economy to be

$$\Lambda_f = \frac{w_f l_f}{I},$$

for each $f \in F$, where w_f and l_f are the price and quantity of factor f . Note that the following equation always holds in equilibrium: $\sum_{f \in F} \Lambda_f = 1$. We refer to the vector of factor income shares by Λ .

Input-output matrix. We stack the expenditure shares of the representative 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 representative agent, we index the household 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

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

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

$$\Psi = \left[\begin{array}{c|ccc|ccc} 1 & \lambda_1 & \cdots & \lambda_N & \Lambda_1 & \cdots & \Lambda_F \\ \hline 0 & \Psi_{11} & \cdots & \Psi_{1N} & \Psi_{1N+1} & \cdots & \Psi_{1N+F} \\ 0 & & \ddots & & & & \\ \hline 0 & \Psi_{N1} & & \Psi_{NN} & \Psi_{NN+1} & \cdots & \Psi_{NN+F} \\ \hline 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & 1 & \vdots \\ \hline 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{array} \right].$$

The first row and column correspond to final demand (good 0). The fact that the first row is equal to the vector of sales shares for goods and factors is a consequence of accounting identities. In an abuse of notation, we refer to the first row of Ψ as λ' (and its transpose as λ). The next N rows and columns correspond to goods, and the last F rows and columns correspond to the factors. Define the $(1 + N + F) \times F$ matrix Ψ^F as the submatrix consisting of the right F columns of Ψ , representing the network-adjusted factor intensities of each good. The sum of network-adjusted factor intensities for every good 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 will use the identities $\lambda' = b'\Psi$ and $\Lambda' = b'\Psi^F$.

4.1 One-Factor Models (Linear PPF)

For intuition, we start by focusing on general equilibrium economies with only one factor of production. Following Proposition 5, when the economy has a single factor of production (or equivalently, a linear PPF), then macro and micro welfare changes are the same. Furthermore, in this case, we can characterize both the change in real GDP and the change in welfare up to a second order.

Proposition 7 (Approximate Macro Welfare vs GDP: Single Factor). *Consider some perturbation in technology, $\Delta \log A$, and final demand, $\Delta \log x$. When the economy has one factor of*

production, the change in real GDP is

$$\begin{aligned} \Delta \log Y \approx & \sum_{i \in N} \lambda_i \Delta \log A_i + \frac{1}{2} \sum_{j \in \{0\} + N} \lambda_j (\theta_j - 1) \text{Var}_{\Omega^{(j)}} \left(\sum_{i \in N} \Psi_{(i)} \Delta \log A_i \right) \\ & + \frac{1}{2} \text{Cov}_{\Omega^{(0)}} \left(\Delta \log x + \left(\sum_{i \in N} \lambda_i \Delta \log A_i \right) \times \varepsilon, \sum_{i \in N} \Psi_{(i)} \Delta \log A_i \right), \end{aligned} \quad (16)$$

where $\text{Cov}_{\Omega^{(j)}}(\cdot)$ is the covariance using the j th row of Ω as the probability weights and $\Psi_{(i)}$ is the i th column of the Leontief inverse. The difference between welfare and GDP is

$$EV^M - \Delta \log Y \approx \frac{1}{2} \text{Cov}_{\Omega^{(0)}} \left(\Delta \log x + \left(\sum_{i \in N} \lambda_i \Delta \log A_i \right) \times \varepsilon, \sum_{i \in N} \Psi_{(i)} \Delta \log A_i \right). \quad (17)$$

Proposition 7 is a general equilibrium counterpart to Proposition 3. We discuss (16) and (17) in turn, starting with (16). The first term in Equation (16) 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 one of these terms captures nonlinearities due to the fact that sales shares can respond to changes in relative prices caused by technology shocks (these effects were emphasized by Baqaee and Farhi, 2019b). The terms on the second line of (16), which are the ones we focus on in this paper, capture changes in sales shares due to changes in preferences or non-homotheticities.

Equation (17) shows that while real GDP correctly accounts for substitution due to supply shocks, it needs to be corrected for substitution due to changes in final demand due to demand shocks or non-homotheticities. Whereas in partial equilibrium, the gap between welfare and real GDP is proportional to the covariance of supply and demand shocks (see Proposition 3), equation (17) shows that in general equilibrium, the relevant statistic is the covariance of demand shocks with a network-adjusted notion of supply shocks not supply shocks per se. Furthermore, Proposition 7 shows that the elasticities of substitution are irrelevant for the gap between welfare and real GDP in one-factor models. This is because relative prices do not change as the equilibrium moves along a linear PPF in response to demand-driven forces. Therefore, demand shocks do not trigger expenditure switching due to the endogenous response of relative prices. When we relax the linearity of the PPF, we see that the elasticities of substitution in production do, in general, affect the gap between welfare and GDP.

We now work through some simple examples to illustrate the intuition in Proposition 7.

Example 1 (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 $\lambda_i = b_i = \Omega_{0i}$ and $\Psi_{(i)}$ is the i th column of the $1 + N + F$ identity matrix $I_{(i)}$. Therefore, Proposition 7 implies

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

Hence, welfare changes are greater than the change in real GDP if productivity and demand shocks are positively correlated. This could happen either because preferences exogenously change to favor high productivity goods, $Cov_b(\Delta \log x, \Delta \log A_i) > 0$, or preferences endogenously change to favor high productivity growth goods due to non-homotheticities, $Cov_b(\varepsilon, \Delta \log A_i) \Delta \log Y > 0$. When shifts in demand are orthogonal to shifts in supply, to a second-order approximation, real GDP measures welfare correctly.

Example 2 (Input-Output Connections). For models with linear PPFs, input-output connections affect the gap between real GDP and welfare in two ways: (1) the impact of technology shocks is bigger when there are input-output linkages because $\Psi_{(i)} \geq I_{(i)}$ and $\lambda_i \geq b_i$; (2) the production network “mixes” the shocks, and this may reduce the correlation of supply and demand shocks by making the technology shocks more uniform. However, since it is the covariance (not the correlation) of the shocks that matters, this means the effects are, at least theoretically ambiguous.

To see these two forces, consider the three economies depicted in Figure 3. Each of these economies has a roundabout structure. Panel 3a depicts a situation where each producer uses only its own output as an input, Panel 3b a situation where all producers use the same basket of goods (denoted by M) as an intermediate input, and Panel 3c a situation where each producer uses the output of the other producer as an input. We compute the correction to GDP necessary to arrive at welfare for each of these cases using Proposition 7. For clarity, we focus on demand shocks caused by instability rather than non-homotheticity, though it should be clear that this does not affect any of the intuitions.

For Panel 3a, assuming that every producer has the same intermediate input share, we get

$$EV^M - \Delta \log Y \approx \frac{1}{1 - \Omega_{ii}} Cov_b(\Delta \log x, \Delta \log A).$$

Hence, as intermediate inputs become more important, the necessary adjustment becomes larger. This is because, for a given vector of preference shocks, the movement in sales shares is now larger due to the roundabout nature of production. On the other hand, for Panel 3b, assuming every producer has the same intermediate input share, we

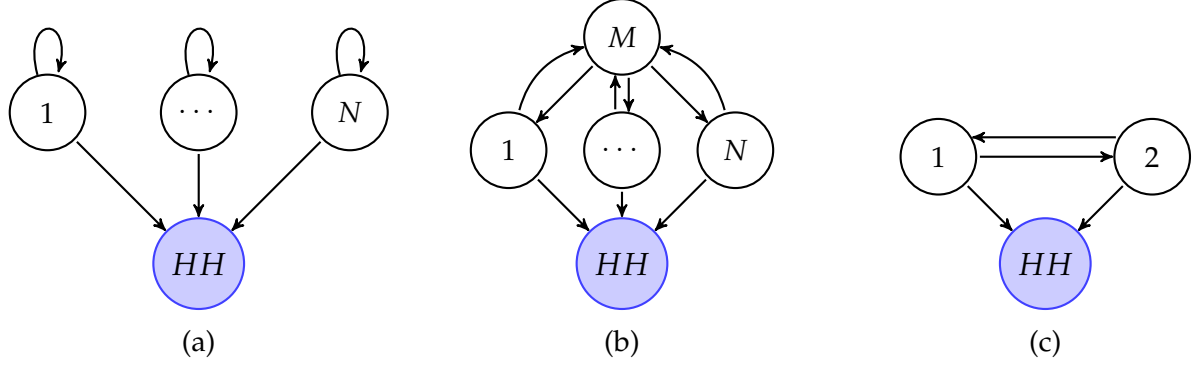


Figure 3: Three different kinds of round-about economy. The arrows represent the flow of goods. The only factor is labor which is not depicted in the diagram.

get

$$EV^M - \Delta \log Y \approx Cov_b(\Delta \log x, \Delta \log A).$$

Hence, in this case, the intermediate input share is irrelevant. Intuitively, since all producers buy the same bundle of materials, a shock to the composition of household demand does not alter the sales of any producer through the supply chain, and hence only the first-round non-network component of the shocks matters. Finally, consider Panel 3c. For clarity, focus on the case where only producer 1 gets a productivity shock ($\Delta \log A_2 = 0$). In this case, the difference between real GDP and welfare is

$$EV^M - \Delta \log Y \approx \frac{1}{1 - \Omega_{12}\Omega_{21}} Cov_b \left(\Delta \log x, \begin{bmatrix} 1 \\ \Omega_{21} \end{bmatrix} \right) \Delta \log A_1.$$

As the intermediate input share Ω_{21} approaches one, the adjustment goes to zero (since the covariance term goes to zero). Intuitively, as Ω_{21} goes to one, the increase in demand for the first producer from a change in preferences is exactly offset by a reduction in demand from the second producer who buys inputs from the first producer. In this limiting case, changes in consumer preferences have no effect on the overall sales share of the first producer.

These three examples serve to illustrate that the effect of input-output networks on the adjustment are theoretically ambiguous but potent.

4.2 Multi-Factor Models (Nonlinear PPF)

To characterize the response of output and welfare to shocks when the PPF is nonlinear we rely on the following result.²⁶

Proposition 8 (Sales and Prices in General Equilibrium). *Consider some perturbation in final demand $d \log x$ and technology $d \log A$. Then changes in prices are, up to a first order²⁷*

$$d \log p = -\Psi d \log A + \Psi^F \left[\frac{\partial \log \Lambda}{\partial \log A} d \log A + \frac{\partial \log \Lambda}{\partial \log x} d \log x \right] \quad (18)$$

Changes in sales shares are

$$\begin{aligned} \lambda_i d \log \lambda_i = & \sum_{j \in \{0\} + N} \lambda_j (\theta_j - 1) \text{Cov}_{\Omega^{(j)}} \left(-d \log p, \Psi_{(i)} \right) \\ & + \text{Cov}_{\Omega^{(0)}} \left(d \log x, \Psi_{(i)} \right) + \text{Cov}_{\Omega^{(0)}} (\varepsilon, \Psi_{(i)}) (\lambda' d \log A), \end{aligned} \quad (19)$$

Changes in factor shares, $\Lambda_f d \log \Lambda_f$ for $f \in F$, are given by the same expression as changes in sales shares.

Unlike the previous propositions, Proposition 8 pins down changes in prices and sales shares only up to a first-order approximation. First-order approximations of changes in prices and sales shares are all we need to plug into the reduced-form expressions in Proposition 2 and Proposition 6. We briefly describe the intuition for (18) and (19).

Equation (18) captures how productivity shocks and changes in factor prices travel downstream through supply chains to affect the marginal cost of downstream produces. Equation (19), on the other hand, captures how changes in relative prices and final demand travel upstream through supply chains to determine the sales of each producer. The first line in (19) corresponds to changes in final demand and intermediate good demand for good i due to changes in relative (goods and factor) prices. The second line of (19) captures instability of final demand (first term on the second line) and non-homotheticity (second term on the second line).

We now work through some simple examples to illustrate how nonlinear PPFs affect our results from earlier.

²⁶Proposition 8 extends Baqaee and Farhi (2019b) to economies where final demand is non-homothetic and unstable.

²⁷In the equations below, $\frac{\partial \log \Lambda}{\partial \log A}$ and $\frac{\partial \log \Lambda}{\partial \log x}$ are $(1 + N + F) \times F$ matrices whose ij th elements are $\frac{\partial \Lambda_j}{\partial \log A_i}$ and $\frac{\partial \Lambda_j}{\partial \log x_i}$, respectively.

Example 3 (Decreasing Returns to Scale). Consider the one-sector model without intermediate inputs in Example 1 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 1. Furthermore, suppose that preferences are homothetic ($\varepsilon_i = 1$ for every i), but potentially unstable ($\Delta \log x \neq 0$). To apply our theorems 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.

We apply Proposition 6 to compute the difference between welfare and real GDP. To do this, we first use Proposition 8 to compute changes in sales shares due to demand shocks:

$$\frac{\partial \lambda_i}{\partial \log x} \cdot d \log x = Cov_{\Omega^{(0)}} \left(d \log x, \Psi_{(i)} \right) + (\theta_0 - 1) Cov_{\Omega^{(0)}} \left(-\frac{d \log p}{d \log x} d \log x, \Psi_{(i)} \right).$$

The factor content of every good $i \in N$ is given by $\Psi_{if}^F = \gamma$ when f is the common factor and $\Psi_{if}^F = 1 - \gamma$ when f is the producer-specific factor. Since the factor shares of each producer are constant, the log change in the producer-specific factor share is the same as the log change in the sales share of that producer $d \log \Lambda_i = d \log \lambda_i$. Therefore, we can replace $d \log p / d \log x$ in the covariance with $(\gamma - 1) (\partial \log \lambda_i / \partial \log x) d \log x$ (the other components of price changes are common to all producers and drop out of the covariance). Plugging this into the expression above and solving yields a closed-form expression for $\partial \lambda_i / \partial \log x$. This allows us to apply Proposition 6 to get the difference between welfare and real GDP up to a second order approximation:

$$EV^M - \Delta \log Y \approx \frac{1}{2} \frac{Cov_{\Omega^{(0)}}(\Delta \log x, \Delta \log A)}{1 + (\theta_0 - 1)(1 - \gamma)}.$$

Note that the denominator disappears when we have constant-returns to scale ($\gamma = 1$) or the elasticity of substitution across goods is one ($\theta_0 = 1$). Outside of these cases, complementarities ($\theta_0 < 1$) amplify the impact 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 4 (Macro vs. Micro Welfare Change). Finally, we demonstrate the difference between macro and micro welfare changes using the previous example. Since the economy in the previous example has multiple factors and unstable preferences, Proposition 5 does not apply. In particular, changes in macroeconomic and microeconomic notions of welfare are different since the PPF is no longer linear.

To illustrate this difference, suppose that only preference shocks are active (there are no supply shocks $\Delta \log A = 0$). By Corollary 2, real GDP changes are zero. Since technology 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 2, micro welfare improves $EV^m > 0$ if preference shocks negatively covary with price changes. By equation (18), changes in prices are

$$d \log p = \Psi^F \frac{\partial \log \Lambda}{\partial \log x} d \log x,$$

with i th element given by $(1 - \gamma) \frac{\partial \log \lambda_i}{\partial \log x} d \log x$. Using the derivations above, we obtain

$$d \log p_i = \frac{(1 - \gamma)}{(1 + (\theta_0 - 1)(1 - \gamma)) \lambda_i} \frac{1}{\lambda_i} \text{Cov}_{\Omega(0)} \left(d \log x, I_{(i)} \right),$$

where $I_{(i)}$ is the i th column of the identity matrix. If there are decreasing returns, $\gamma < 1$, then a positive demand shock for i raises the price of i . The change in the price is amplified if goods are complements and mitigated if goods are substitutes. We can now apply Proposition 3 to obtain the change in micro welfare, up to a second order,

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

With decreasing returns to scale ($\gamma < 1$), micro welfare decreases since 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.

On the other hand, when the economy has increasing returns to scale ($\gamma > 1$), micro welfare increases in response to demand shocks. Intuitively, in this case, increased de-

mand for a good lowers the price of that good, which makes the consumer better off.²⁸ Of course, from a societal perspective, welfare has not changed, since the production possibility set of the economy has not changed.

5 Applications

In this section, we consider three applications of our results. The first application is to the problem of long-run growth and the difference between welfare-relevant and measured aggregate productivity growth in the presence of non-homotheticities in demand. The second application shows that correlated firm-level supply and demand shocks drive a wedge between measured real GDP and welfare even in the short-run. Our last application considers how demand instability can make measured real GDP an unreliable metric for changes in production, and we illustrate this point for the Covid-19 crisis.

5.1 Long-Run Growth and Structural Transformation

Baumol (1967) showed that, as economies grow, sectors with lower relative productivity growth rates expand (in terms of sales and value-added) relative 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 oftentimes called Baumol's cost disease.

The causes of changes in industry mix over time, called structural transformation, are the subject of a large literature. As discussed by Buera and Kaboski (2009) and Herrendorf et al. (2013), there are two primary reasons the literature offers for why economies undergo structural transformation. The first reason is complementarities, either in demand or in production, that mean that industries with faster productivity growth, whose relative prices fall over time, shrink as a share of overall expenditures. The second reason is non-homotheticities or instability in final demand, whereby marginal propensities to consume across different sectors change as the economy grows, resulting in more expenditures on sectors with lower productivity growth (i.e. services).

Either way, changes in sales shares will affect aggregate productivity growth. Following Nordhaus et al. (2008), aggregate productivity growth can be decomposed into two

²⁸If the economy has increasing returns to scale, then the decentralized equilibrium is potentially inefficient. However, the propositions regarding micro welfare changes, which take changes in prices as given, do not require that the decentralized equilibrium be efficient.

terms:

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

where $\lambda_{i,t}$ is the sales shares of industry i in period t and $\Delta \log A_{i,t}$ is the growth in gross-output productivity over that time period.²⁹ The first term captures changes in aggregate TFP if industry-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 4 implies that, for the purposes of welfare, changes in sales shares due to non-homotheticities or demand instability must be treated differently to changes in sales shares due to complementarities. In particular, the welfare-relevant measure of the change in TFP is

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

where λ^{ev} is the hypothetical sales-shares of each industry holding fixed final preferences and income-level — that is, sales shares after they have been purged from changes due to factors other than changes in relative prices.

Two polar extremes. Computing these terms requires an explicit structural model of the economy. However, there are two polar cases in which the TFP adjustment term can be calculated without specifying the detailed model. On the one hand, demand is stable and homothetic, and changes in sales shares are due only to relative price changes (complementarities). On the other hand, there are no complementarities (as in a Cobb-Douglas economy) and changes in sales shares are only due to non-homotheticity or demand instability. If structural transformation is driven by a combination of complementarities and demand instability or non-homotheticities, then the change in welfare TFP will be somewhere in between these two cases. The following corollary of Proposition 4 summarizes the change in welfare-TFP in these two polar cases.

²⁹Technically, this is an approximation, since we define aggregate TFP in continuous time but the data is 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 $t + 1$ averages.

Corollary 3. *If changes in sales shares are due only due only to complementarities, then*

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

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

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

In the first case, since preferences are homothetic and stable, welfare-TFP is equal to TFP in the data. In the second case, since there are no complementarities in production or demand, sales shares do not respond to productivity changes. In order to hold utility and preferences fixed at their final value, we must compute welfare-TFP using terminal sales shares.

To quantify Corollary 3, we use US-KLEMS data on sales shares and TFP growth for 61 private-sector industries. We calculate changes in industry-level gross-output TFP following the methodology of Jorgenson et al. (2005) and Carvalho and Gabaix (2013). 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 (or 60%). Hence, Baumol’s cost-disease caused aggregate TFP to fall by -10 log points, reducing aggregate productivity growth by around 23 percent (from 78% to 60%). Using Corollary 3, if we assume that structural transformation is due solely to non-homotheticities or demand instability, then the growth in welfare-relevant TFP from 1948-2014 has been 37 log points (or 46%) instead of the measured 47 log points (or 60%) — that is, to say, a 23 percent additional reduction in the growth rate. In other words, the adjustment necessary to correct for demand-driven structural transformation is roughly twice as big (in log terms) as the one necessary to account for complementarities related to relative prices.

To understand why the necessary adjustment is roughly twice as big, consider the second-order approximation in Proposition 4. Up to a second-order approximation, we can write

$$\Delta \log TFP^{welfare} \approx \Delta \log TFP + \frac{1}{2} \left[\sum_{i \in N} \frac{\partial \lambda_i}{\partial \log x} \Delta \log x + \frac{\partial \lambda_i}{\partial \log v} \Delta \log v \right] \Delta \log A_i, \quad (20)$$

where

$$\Delta \log TFP \approx \sum_{i \in N} \lambda_{i,t_0} \Delta \log A_i + \frac{1}{2} \sum_{i \in N} \Delta \lambda_i \Delta \log A_i.$$

If changes in sales shares are due entirely to demand-driven factors, then the term in square brackets in (20) is equal to $\sum_{i \in N} \Delta \lambda_i \Delta \log A_i$, so

$$\Delta \log TFP^{\text{welfare}} \approx \sum_{i \in N} \lambda_{i,t_0} \Delta \log A_i + \sum_{i \in N} \Delta \lambda_i \Delta \log A_i.$$

In other words, the adjustment to the initial sales shares must be roughly twice as large as the adjustment to the initial sales shares caused by complementarities. These second-order approximations are more accurate if changes in sales shares are well-approximated by linear time trends, and the surprising accuracy of the second-order approximation is a result of this fact.

The reasoning above implies that the welfare adjustment to bring measured TFP in line with welfare may be as high as 23 percent. However, it may also be as small as zero, if the data can be perfectly explained using only complementarities.

Quantitative illustration away from two polar extremes. In practice, both complementarities and non-homotheticities are likely to play an important role in explaining structural transformation. To dig deeper into the size of the welfare adjustment, we use a simplified version of the model introduced in Section 4 calibrated to the US economy, accounting for input-output linkages and complementarities, and use the model to quantify the size of the welfare-adjustment as a function of the elasticities of substitution.

Remarkably, Proposition 4 implies that to compute the welfare-relevant change in TFP, we must only supply the information necessary to compute λ^{ev} — that is, we do not need to model the non-homotheticities or demand-shocks themselves, and the exercise requires no information on the functional form of non-homotheticities or the slope of Engel curves or magnitude of income elasticities *conditional* on knowing the elasticities of substitution.

We map the model to the data as follows. We assume that the constant-utility final demand aggregator has a nested-CES form. There is an elasticity θ_0 across the three groups of industries: primary, manufacturing, and service industries. The inner nest has elasticity of substitution θ_1 across industries within primary (2 industries), manufacturing (24 industries), and services (35 industries).³⁰ Production functions are also assumed to have

³⁰In order to map this nested structure to our baseline model, good 0 is a composite of good 1-3, where good 1 is a composite of primary industries, good 2 is a composite of manufacturing industries, and good

nested-CES forms: there is an elasticity of substitution θ_2 between the bundle of intermediates and value-added, and an elasticity of substitution θ_3 across different types of intermediate inputs. For simplicity, we assume there is only one primary factor of production (a composite of capital and labor). We solve the non-linear model by repeated application of Proposition 8.³¹

We calibrate the CES share parameters so that the model matches the 2015 input-output tables provided by the BEA. For different values of the elasticities of substitution $(\theta_0, \theta_1, \theta_2, \theta_3)$ we feed changes in industry-level TFP (going backwards, from 2015 to 1960) into the model and compute the resulting change in aggregate TFP. This number represents the welfare-relevant change in aggregate TFP. We report the results in Table 1.

Table 1: The percentage change in measured and welfare-relevant TFP in the US from 1960 to 2005.

$(\theta_0, \theta_1, \theta_2, \theta_3)$	(1,1,1,1)	(0.5,1,1,1)	(1,0.5,1,1)	(1,1,0.5,1)	(1,1,1,0.5)
Welfare TFP	46%	46%	54%	48%	55%
Measured TFP	60%	60%	60%	60%	60%

The first column in Table 1 shows the change in welfare-relevant TFP assuming that there are no complementarities (all production and consumption functions are Cobb-Douglas). In this case, all changes in sales shares in the data are driven by non-homotheticities or demand-instability, and hence welfare-relevant TFP has grown more slowly than measured TFP, exactly as discussed in the previous section. The other columns show how the results change given lower elasticities of substitution. As we increase the strength of complementarities, we reduce the size of changes in sales shares due to non-homotheticities or demand-instability, which in turn reduces the gap between measured and welfare-relevant productivity growth.

Table 1 also shows that not all elasticities of substitution are equally important. The results are much more sensitive to changes in the elasticity of substitution across more disaggregated categories, like materials, than aggregated categories, like agriculture, manufacturing, and services. In Appendix H, we show that this is an intuitive consequence of Propositions 5 and 7.

3 is a composite of service industries. Goods 4-65 are the disaggregated industries. Finally, good 66 is the single factor of production.

³¹We calculate TFP shocks by industry in the data allowing for cross-industry variation in capital and labor shares, and then feed these shocks as a primitive into our one-factor model (alternatively, in order to allow for industry variation in capital intensity while keeping the one-factor simplicity, we must assume that labor and capital are perfect substitutes).

5.2 Aggregation Bias with Firm-Level Shocks

In the previous application, we considered a long-run industry-level application. Since industry-level sales shares are relatively stable over short-horizons, given industry-level data, biases in real GDP and consumption are likely to be modest at high frequency. However, this does not mean that these biases are necessarily absent from 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, even if all firms are infinitesimal within their industry. Furthermore, these biases in industry-level data are not diversified away as we aggregate, even if all industries are infinitesimal in the aggregate economy. If these firm-level supply and demand shocks are persistent, then this bias becomes even larger as we lengthen the horizon. Finally, if we start with industry-level (rather than firm-level) data, we are ruling out the existence of these biases by construction.

To make these points formally, we rely on a specification of our model with an explicit industrial structure.

Definition 6 (Industrial Structure). An economy has an *industry structure* if the following conditions hold:

- i. Each firm i belongs to one, and only one, industry I . Firms in the same industry share the same constant-returns-to-scale production function up to a firm-specific Hicks-neutral productivity shifter A_i .
- ii. The representative household has homothetic preferences over industry-level goods, where the I th industry-level consumption aggregator is

$$c_I = \left(\sum_{i \in I} \bar{b}_{iI} x_i c_i^{\frac{\zeta_I - 1}{\zeta_I}} \right)^{\frac{\zeta_I}{\zeta_I - 1}},$$

where c_i are consumption goods purchased by the household from firm i in industry I and x_i are firm-level demand shocks.

- iii. Inputs purchased by any firm j from firms i in industry I are aggregated according to

$$x_{jI} = \left(\sum_{i \in I} \bar{s}_{iI} x_{ji}^{\frac{\sigma_I - 1}{\sigma_I}} \right)^{\frac{\sigma_I}{\sigma_I - 1}},$$

where x_{ji} are inputs purchased by firm j from firm i , and \bar{s}_{iI} is a constant.

Input-output and production network models that are disciplined by industry-level data typically have an industry structure of the form defined above. For such economies, we can characterize the bias in real GDP relative to welfare easily.

Proposition 9 (Aggregation Bias). *For models with an industry structure, in response to firm-level supply shocks $\Delta \log A$ and demand shocks $\Delta \log x$, we have*

$$\Delta \log EV^M \approx \Delta \log Y + \frac{1}{2} \sum_I b_I \text{Cov}_{b_{(I)}}(\Delta \log x_{(i)}, \Delta \log A_{(i)}) + \Theta,$$

where b_I is industry I 's share of final demand and $b_{(I)}$ is a vector whose i th element is b_i/b_I if i belongs to industry I and zero otherwise. The scalar Θ is the gap between real GDP and welfare in a version of the model with only industry-level shocks.³²

In words, Proposition 9 implies that if firms' productivity and demand shocks are correlated with each other (but not necessarily across firms), then there is a gap between real GDP and welfare that does not appear in an industry-level specification of the model. Furthermore, this bias is, to a second-order, additive. That is, the overall bias is the sum of the industry-level bias (that we studied in the previous section) plus the additional bias driven by within-industry covariance of supply and demand shocks.

Identifying supply and demand shocks at the firm-level is notoriously difficult, and many papers impose that these shocks must be orthogonal in order to estimate them (or to estimate demand elasticities). However, a correlation between firm-level supply and demand shocks is a natural prediction in models with endogenous customer acquisition (Arkolakis, 2016; Foster et al., 2016; Fitzgerald et al.; 2016), as we discuss in Section 6. A detailed quantitative investigation of this issue is beyond the scope of this paper, but Proposition 9 can be used for a back-of-the-envelope calculation.

As a numerical illustration, suppose that firm- or product-level demand shocks and TFP shocks at annual frequency have a standard deviation of 0.5 and 0.2 log points, and suppose that they have a correlation of 0.2. Proposition 9 implies that, up to a second-order approximation and at annual frequency, the gap between welfare-relevant industry

³²More specifically,

$$\Theta = \frac{1}{2} \sum_I \sum_j \left[\mathbb{E}_{b_I}(\Delta \log A_{iI}) \left(\frac{\partial b_I}{\partial \log x_j} \right) + \mathbb{E}_{s_I}(\Delta \log A_{iI}) \left(\frac{\partial \lambda_I^f}{\partial \log x_j} \right) \right] \Delta \log x_j,$$

where $\lambda_I^f = \lambda_I - b_I$ is the sales share of industry I excluding sales to the household. In words, Θ is proportional to the product of changes in industry-level sales shares caused by demand shocks times the industry-level (sales-weighted) productivity shocks.

output and measured industry output is around $\frac{1}{2} \times 0.2 \times 0.5 \times 0.2 = 1\%$.³³ If these industry-level covariances are the same across all industries, then the gap between real GDP and welfare induced by firm-level shocks is also 1% at annual frequency. This is a one-off steady-state adjustment that does not compound over time (every year, the change in welfare would be 1% higher than real GDP) if the shocks are independent over time.

To continue this thought experiment, note that if supply and demand shocks at the firm level are persistent, then the bias grows over time. For example, suppose that firm-level supply and demand shocks are both AR(1) processes with persistence parameters $\rho^A \in [0, 1]$ and $\rho^x \in [0, 1]$. Then the gap between welfare and real GDP in period T relative to some period 0 due to firm-level shocks is (suppressing industry sub-indices)

$$Cov_b(\Delta \log x_T, \Delta \log A_T) = \frac{1 - (\rho_x \rho_A)^T}{1 - (\rho_x \rho_A)} Cov_b(\Delta \log x_1, \Delta \log A_1).$$

Hence, as we expand the horizon, the gap between real GDP and welfare becomes larger. In the limit where supply and demand shocks are correlated random walks, the gap expands without bound as we lengthen the horizon. If the persistence parameters are both equal to 0.9, then the asymptotic gap between welfare and real GDP is around 5 times larger than the one-period gap.

5.3 Path Dependence of Real GDP with an Application to Covid-19

In our last application, we switch from focusing on welfare to considering the properties of real GDP in the presence of unstable or non-homothetic preferences. We apply our analysis to one specific and noteworthy instance of demand instability: the onset of the Covid-19 crisis. As shown by Corollary 1 changes in real consumption measure changes in welfare accurately if, and only if, preferences are stable and homothetic. The question that begs to be asked is: *does the change in real GDP or real consumption measure the change in any stable function of allocations when preferences are unstable or non-homothetic?* The gradient theorem for line integrals provides an answer. If sales shares are not derivatives of some stable scalar-valued function, then the change in real GDP between two allocations can be any arbitrary number depending on the path taken (see Hulten, 1973).

This means that whenever preferences are non-homothetic or unstable, changes in real consumption (or real GDP) generically depend on the path and timing of shocks, unlike

³³The gap would be around 2% based on statistics for TFPQ and demand shocks reported in Table 2 of Eslava and Haltiwanger (2021) using Colombian data. We obtain very similar numbers if we compute the gap exactly (numerically) or if we use the second-order approximation in Proposition 9

changes in welfare, which depend only on the initial and final allocation.³⁴ To see this, suppose final demand is homothetic but unstable, and the economy faces supply shocks $\Delta \log A$ and demand shocks $\Delta \log x$. Consider two different paths for how the shocks unfold: under path *SD* the supply shocks arrive first and then the demand shocks, and under *DS* the demand shocks arrive first and then the supply shocks. A corollary of Proposition 6 is then the following.

Corollary 4 (Path Dependence and Timing of Shocks). *If preferences are homothetic, then*

$$\Delta \log Y^{SD} = \log \frac{e(p_{t_0}, v(A_{t_1}; x_{t_0}); x_{t_0})}{e(p_{t_0}, v(A_{t_0}, x_{t_0}); x_{t_0})}, \quad \Delta \log Y^{DS} = \log \frac{e(p_{t_0}, v(A_{t_1}; x_{t_1}); x_{t_1})}{e(p_{t_0}, v(A, x_{t_1}); x_{t_1})},$$

where $\Delta \log Y^{SD}$ and $\Delta \log Y^{DS}$ denote real GDP compute under *SD* and *DS*, respectively.

That is, the change in real GDP under the first path measures something different than the change in real GDP under the second path. In general, the increase in real GDP is larger if price falls (or productivity rises) later for goods that are subject to earlier positive demand shocks.³⁵ If we go from the initial allocation to the final one along the path *SD* and then come back along *DS*, real GDP will be different even though we are back at the exact same allocation. Hence, how fast the economy recovers from a shock like the Covid-19 crisis can be affected by irrelevant details like the order in which supply and demand conditions revert to normal.

Disconcerting as this path dependence problem of real GDP (and other Divisia indices) may be, it is tempting to dismiss it as a theoretical curiosity. However, far from being an exotic theoretical possibility, path-dependence is a well-known problem for practitioners when dealing with high-frequency data. Most manuals on index number construction, for example OECD et al. (2004), warn against using chain-weighted indices for high frequency data.

The term “*chain drift*” bias in national accounting describes a situation when a chained index measures an overall change between t_0 and t_1 even though all prices and quantities in t_0 and t_1 are identical. The gradient theorem for line integrals implies that this must happen if the weights are *not* derivatives of a stable potential function, even if prices and quantities of all goods are well-measured over time.

Chain drift bias appears when movements in prices and quantities are oscillatory,

³⁴If real consumption is calculated using fixed weights over time rather than as a Divisia index, then changes in real consumption are not path dependent, but differ from welfare changes (even if preferences are stable and homothetic).

³⁵This implies that, in the presence of demand instability, the order of reforms that increase sectoral productivity matters for the magnitude of the overall increase in real GDP

where changes that take place over some periods are reversed in subsequent periods. This happens frequently at sub-annual frequencies, so the usual advice given to practitioners is to avoid these problems by not chaining their data too frequently. Intuitively, if we only chain indices at annual or lower frequencies, then it will very rarely be the case that prices and quantities will revert back to their previous values, and hence, we will not observe changes in the aggregate even though every price and quantity has returned to an initial value.

Chaining at low frequency does not solve the problem however, it simply avoids having to confront it by ensuring that prices and quantities never revert back to their previous values. But the path dependence problem that causes chain drift does not become less severe at lower frequencies, it simply becomes less noticeable. The approach we advocate in this paper, which corrects the index for instability and non-homotheticity, has the virtue of restoring path independence.

The Covid-19 recession provides an illustrative case in point. During the first two quarters of 2020, employment and production rapidly contracted in most economies around the world. This contraction was unusual in that it fell very unevenly across different industries. The sales shares of some sectors, like air transportation or entertainment venues, collapsed while those of other sectors, like retail trade, rose. These changes are difficult to justify on the basis of observed price changes since relative price movements across sectors were relatively modest. Given the fact that producing and consuming some goods requires much more face-to-face contact than others, it is natural to suppose that non-price factors likely played an important role in the observed changes in sales shares.³⁶ That is, final demand across sectors as a function of measured prices was unstable during this period.

We now consider how sensitive the change in real GDP can be to the exact timing of the supply and demand shocks. To answer this question, we use a modified version of the structural model introduced above. Since we are interested in 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 then calibrate share parameters to match the

³⁶These non-price factors could be changes in quality, changes in preferences, or changes in government policies during the pandemic. For our purposes in this application (i.e. path dependence of real GDP as measured in the data), it does not matter what is the source of demand instability at measured prices. Of course, in order to assess the welfare costs of Covid-19, the source of demand instability is important. However, since consumers do not make choices about whether or not they live in a world with Covid-19, the welfare costs of Covid-19 are not revealed by their choices. For example, choices between going to restaurants or cooking at home reveal preferences between these two goods conditional on living in a world with a pandemic. What is not revealed by these choices are preferences over restaurants with and without pandemics, and cooking at home with and without pandemics. For this reason, we refrain from analyzing the welfare costs of Covid-19 in this application.

input-output table in 2018, and consider a range of elasticities of substitution. Following Baqaee and Farhi (2020), we hit the economy with a vector of primitive supply and demand shocks.³⁷ The supply shocks are reductions in factor endowments, calibrated to match peak-to-trough reductions in hours worked by sector from January, 2020 to May, 2020. The primitive demand shocks are shifters in final demand, calibrated to match the observed peak-to-trough reductions in personal consumption expenditures by sector from January, 2020 to May, 2020 (given the supply shocks). The details of this calibration exercise can be found in Appendix D.

Table 2 reports the results. We consider three different calibrations informed by empirical estimates from Atalay (2017) and Boehm et al. (2015): high complementarities, medium complementarities, and no complementarities (Cobb-Douglas). The high complementarity scenario sets the elasticity of substitution across consumption goods (outer and inner nests) 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 2: The change in real GDP under alternative paths for the arrival of supply and demand shocks between February 2020 to May 2020.

Elasticities	High compl.	Medium compl.	Cobb-Douglas
Supply then demand	-16.2%	-12.5%	-10.8%
Demand then supply	-10.1%	-9.4%	-9.0%
Simultaneous supply and demand	-12.1%	-10.6%	-9.8%

Table 2 shows that the reduction in real GDP is much larger when supply shocks arrive first before the demand shocks, even though, the difference in welfare and the difference in physical allocations in every scenario is the same. This is because changes in demand are positively correlated with changes in supply shocks. That is, the reduction in the output of a sector reduces GDP by less if the households have already reduced their spending on that sector's output. Hence, reductions in the supply of transportation services or food and beverage stores are less costly if demand for those sectors contracts first, than if production falls followed by a reduction in demand.

³⁷For related analysis of Covid-19 induced supply shocks, see e.g. Bonadio et al. (2020) and Barrot et al. (2020). For related analysis of Covid-19 induced demand shocks, see Cakmakli et al. (2020).

The wide gaps we document here suggest that if, during the recovery, the supply and demand shocks do not disappear in exactly the same way as they arrived, measured real GDP can be higher or lower than it was before the crisis, even if the economy returns exactly to its pre-Covid allocation. And this difference can be as much as 6% — equivalent to several years worth of economic growth. Hence, during episodes where final demand is unstable, real GDP becomes an increasingly unreliable guide for measuring output.

6 Extensions

In this section, we briefly summarize how our results can be extended in different ways.

Extensive margin. If preference instability or non-homotheticity causes a consumer to begin purchasing a good in t_1 that she did not consume in t_0 (or to stop consuming a good that she was previously consuming), then our global and local formulas apply to that consumer without change.

To make this more explicit, consider a consumer whose preferences are represented by the utility function

$$\left(\int_0^{x^*} c(z)^{\frac{\sigma-1}{\sigma}} dz \right)^{\frac{\sigma}{\sigma-1}}, \quad (21)$$

where goods are indexed by $z \in [0, 1]$ but the consumer only values goods $z \in [0, x^*]$. In this situation, x^* is a preference parameter, where goods $z \in (x^*, 1]$ are available at a finite price, but the consumer chooses not to consume them.

Consider how the welfare of the consumer changes accounting for the fact that x^* can change between t_0 and t_1 . The following is an application of Lemma 1.

Proposition 10 (New Goods Due to Taste Shocks). *Consider a household with preferences defined by (21). Up to a second-order approximation,*

$$\Delta \log EV^m \approx \Delta \log Y + \frac{1}{2} b(x^*) \Delta x^* [\mathbb{E}_b [\Delta \log p] - \Delta \log p(x^*)].$$

In words, the gap between welfare and real GDP depends on product of sales shares at the cut-off $b(x^*)$, the change in the cut-off Δx^* , and the difference between inflation at the cut-off versus average inflation. If new goods are added $\Delta x^* > 0$, and the new goods experienced lower than average inflation, then welfare is higher than what is detected by real GDP. However, this adjustment is second-order (since it involves products of Δ), and to a first-order, real GDP is equal to the true change in the cost of living.

It is interesting to contrast Proposition 10 to the well-known new-goods adjustment due to Feenstra (1994), which, to a first-order approximation, is

$$\Delta \log EV^m = \frac{1}{1-\sigma} \Delta \log \left[\int_{\mathcal{C}} b(z) dz \right] - \mathbb{E}_b [\Delta \log p], \quad (22)$$

where \mathcal{C} is the set of continuing goods and the integral is their share in expenditures. The difference in these results is due to a difference in interpretation. Under the interpretation in Proposition 10, the change in the extensive margin is caused by a change in tastes — that is, the goods were previously available to the consumer in the initial period but the consumer chose not to consume them (or goods are available in the final period, but the consumer chose to stop consuming them). Therefore, when we calculate welfare changes, we simply need to adjust the price index so that it accounts for the price of goods that the consumer is choosing to consume in the final period. On the other hand, under (22), when we compute the change in welfare, we assume that the consumer is unable to consume the new goods in the past or can no longer consume the disappearing goods in the present. That is, under (22), when goods are not consumed they are valued by the consumer but the implicit price is infinity.

Therefore, if a good is available in t_0 and t_1 , but the consumer does not consume the good in period t_0 and does consume the good in t_1 (due to, for example, advertising), an application of (22) is not innocuous. If the change in consumer behavior is due to a change in tastes, as opposed to a change in availability, then no adjustment is necessary to a first-order, and to a second-order, the relevant adjustment is the one in Proposition 10.

Endogenous separable arguments in the utility function (e.g. leisure). If there are goods in the utility function that are endogenously chosen but not measured, then an all-encompassing welfare measure must impute shadow prices for these goods (see Jones and Klenow, 2016). For concreteness, suppose that leisure is the non-measured argument in the utility function. If leisure is separable from market goods, so that preferences over c are stable when the quantity of leisure changes, then our baseline results apply to the measured-good component of welfare, even if leisure changes.

Endogenous non-separable arguments in the utility function (e.g. advertising). If the parameters of the utility function x are not separable from goods c , then our welfare questions ask how changes in constraints over c affect welfare holding fixed x . That is, we do not attempt to answer how a change in x itself affects welfare, which may or may not be

a question that can be answered. A salient example of x can be advertising, which can change households ranking over different consumption bundles, and is obviously non-separable from market goods. In principle, advertising may have value to the consumer — that is, the consumer can have preferences over the amount of advertising they wish to be exposed to. If advertising is informative, the consumer’s utility may be increasing in advertising, and equally plausibly, if advertising is manipulative, then utility may be decreasing in advertising. We do not attempt to answer the question of how much the household values advertising, instead, we hold fixed the amount of advertising³⁸ (or indeed the weather, chemicals in the brain, and whatever else that affects valuations over consumption bundles), and measure how changes in the availability of market goods affects welfare.

Interestingly, unlike random fluctuations in tastes, advertising is a purposeful economic activity, and therefore, models of advertising and consumer acquisition, for example Arkolakis (2016), explicitly predict 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. This positive correlation means that we should expect real GDP or real consumption measures to be systematically biased in situations where advertising plays a large role in consumption choices.

Beyond CES. Our results in Section 4 can be generalized beyond CES functional forms relatively easily. In Appendix C, we discuss how Proposition 8 must be adjusted to allow for non-CES production and utility functions.

Wedges in final demand. Our analysis throughout the paper assumes optimization. That is, either the consumer is optimizing given a budget constraint or society is optimizing given a production possibility set. To consider situations where consumers or society is not optimizing, we can introduce unobserved wedges in final demand which distort consumers’ choices. Such wedges would allow us to compare changes in welfare, given fixed preferences, caused by a change in allocations, rather than changes in welfare caused by a change in constraints. For more details on this approach, see Appendix B.

Heterogeneous agents. Our microeconomic welfare results can be applied in economies without representative agents (including open economies). Our macroeconomic welfare

³⁸In this sense, our approach is related to Dixit and Norman (1978), who study the welfare implications of advertising at either pre- or post-advertising preferences. As argued by Fisher and McGowan (1979), this does not answer the question of what is the value of advertisement taking into account the change in tastes.

results relied on the assumption that there is a representative agent because, in this case, the competitive equilibrium is maximizing this agent's welfare.

In economies with heterogeneous agents and non-aggregable preferences, the social welfare function (SWF) is not unambiguous. In such economies, every competitive equilibrium maximizes a potentially different SWF (see Negishi, 1960). Whereas to a first-order, changes in real output will coincide with changes in the SWF that is being maximized in the initial competitive equilibrium, at higher orders this relation breaks down. However, if commits to a particular SWF that is maximized in the initial equilibrium, results like the ones we developed in Section 3 can be used to correct the gap between social welfare and real GDP.

For example, consider the following utilitarian SWF which weights the welfare of each agent i by their initial share in expenditures,

$$W = \sum_i \frac{\bar{p}_i \bar{Y}_i}{\sum_j \bar{p}_j \bar{Y}_j} \left[\frac{EV_i^m}{E\bar{V}_i^m} \right]$$

This SWF is maximized in the initial equilibrium. Since the parameters of this SWF change in response to shocks that redistribute income, it is as if the social planner has unstable preferences. This instability can create a gap between social welfare and GDP even if preferences of individual agents are stable and homothetic.

For this case, in Appendix E we show that under the assumptions in Section 3, to a second-order approximation,

$$\Delta \log W = \Delta \log Y + \frac{1}{2} \text{Cov}_\chi (\mathbb{E}_b[\Delta \log p], \Delta \log I - \mathbb{E}_b[\Delta \log p]),$$

where χ is the initial distribution of expenditures by each agent, and $\mathbb{E}_b[\Delta \log p]$ is the vector of inflation rates and $\Delta \log I$ is the vector of nominal income changes for each household. The change in social welfare is greater than the change in real GDP if changes in real income covary positively with inflation across households. There is no bias if preferences are aggregable, in which case $\Delta \log p$ is uniform across households, if real income growth is uniform or, more generally, if changes in household-level inflation rates are uncorrelated with changes in real income.³⁹

³⁹As usual, notice that since the terms in the adjustment $\mathbb{E}_b[\Delta \log p]$ and $\Delta \log I - \mathbb{E}_b[\Delta \log p]$ only need to be approximated to a first-order, since they are multiplied, and we are only keeping second-order terms.

7 Conclusion

In this paper, we provide a characterization of the gap, due to substitution bias, between standard measures of consumption and welfare that appear when preferences are non-homothetic or unstable. We characterize this bias in both partial and general equilibrium, and show that it can be large over long horizons relevant for long-run growth, and short-horizons when expenditure shares change rapidly (as in they do with firm or product-level sales shares).

Although our motivation and applications have focused on shocks across time, our results can also be applied to compare welfare across locations in space. Variation in tastes is likely to be even more significant across space than across time (see, e.g. Deaton, 2003 and Argente et al., 2020). Applying our results in a spatial context is an interesting avenue for future work.

References

- Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica* 80(5), 1977–2016.
- 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. Müller (2019). A theory of structural change that can fit the data.
- Argente, D. O., C.-T. Hsieh, and M. Lee (2020, September). Measuring the cost of living in Mexico and the US. Working Paper 27806, National Bureau of Economic Research.
- 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* (Forthcoming).
- 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. Lein, and J. Vogel (2021). Unequal expenditure switching: Evidence from Switzerland.
- 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 E. Farhi (2019a). Networks, barriers, and trade. Technical report, National Bureau of Economic Research.
- Baqae, D. and E. Farhi (2020). Supply and demand in disaggregated Keynesian economies with an application to the COVID-19 crisis. Technical report, National Bureau of Economic Research.

- Baqae, D. R. and E. Farhi (2019b). The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem. *Econometrica* 87(4), 1155–1203.
- Barrot, J.-N., B. Grassi, and J. Sauvagnat (2020). Sectoral effects of social distancing. *Available at SSRN*.
- Basu, S., L. Pascali, F. Schiantarelli, and L. Serven (2012). Productivity and the welfare of nations. Technical report, National Bureau of Economic Research.
- Baumol, W. J. (1967). Macroeconomics of unbalanced growth: the anatomy of urban crisis. *The American economic review*, 415–426.
- Boehm, C., A. Flaaen, and N. Pandalai-Nayar (2015). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 tōhoku earthquake.
- Bonadio, B., Z. Huo, A. A. Levchenko, and N. Pandalai-Nayar (2020). Global supply chains in the pandemic. Technical report.
- 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.
- 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 (2018). Production networks: A primer.
- 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: Journal of the Econometric Society*, 1393–1414.
- Comin, D. A., D. Lashkari, and M. Mestieri (2015). Structural change with long-run income and price effects. Technical report, National Bureau of Economic Research.
- Deaton, A. (2003). Adjusted indian poverty estimates for 1999-2000. *Economic and political Weekly*, 322–326.
- Dixit, A. and V. Norman (1978). Advertising and welfare. *The Bell Journal of Economics*, 1–17.
- Domar, E. D. (1961). On the measurement of technological change. *The Economic Journal* 71(284), 709–729.
- Eslava, M. and J. C. Haltiwanger (2021). The life-cycle growth of plants: The role of productivity, demand and wedges. Technical report, National Bureau of Economic Research.
- Fally, T. (2020). Integrability, generalized separability, and a new class of demand systems.
- Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *The American Economic Review*, 157–177.
- Fisher, F. M. and J. J. McGowan (1979, Autumn). Advertising and Welfare: Comment. *Bell Journal of Economics* 10(2), 726–727.
- Fisher, F. M. and K. Shell (1968). *Taste and quality change in the pure theory of the true cost of living index*. Chicago: Aldine Publishing.
- Fitzgerald, D., S. Haller, and Y. Yedid-Levi (2016). How exporters grow. Technical report, National Bureau of Economic Research.
- Foster, L., J. Haltiwanger, and C. Syverson (2016). The slow growth of new plants: learning about demand?

- Economica* 83(329), 91–129.
- Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica* 79(3), 733–772.
- Hamilton, B. W. (2001, June). Using engel’s law to estimate cpi bias. *American Economic Review* 91(3), 619–630.
- 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 transformation. *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.
- 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.
- Negishi, T. (1960). Welfare economics and existence of an equilibrium for a competitive economy. *Metroeconomica* 12(2-3), 92–97.
- 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. et al. (2008). Baumol’s diseases: A macroeconomic perspective. *The BE Journal of Macroeconomics* 8(1), 1–39.
- OECD, T. W. Bank, T. U. N. E. C. for Europe, S. O. of the European Communities, and Luxembourg (2004). *Consumer price index manual: Theory and practice*. International Monetary Fund.
- 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.

[Appendix To Be Added]