Demand Shocks and Real GDP Measurement

Emi Nakamura *
Harvard University

June 27, 2004

Preliminary and Incomplete

Comments welcome

Abstract

This paper investigates the implications of demand shocks for real GDP measurement. Traditionally, index number theory proceeds under the assumption of fixed preferences. However, historical data on alternative quantity indexes for US real GDP suggest large deviations from the standard theoretical predictions. This paper suggests explanations for this puzzle based on the role of demand shocks in generating economic fluctuations, and investigates the implications of demand shocks for utility-based interpretations of real GDP.

Keywords: Consumer substitution bias, real GDP measurement

JEL Classifications: C43, O47, E23.

*I would like to thank W. Erwin Diewert for extremely helpful and inspiring comments on this paper. I would also like to thank Susanto Basu, Dan Benjamin, N.G. Mankiw, Robert Parker, Marc Prud'homme, Ricardo Reis, Julio Rotemberg, Jesse Shapiro, Assaf Ben Shoham, Martin Weitzman, seminar participants at Harvard University and Statistics Canada, and particularly Alice Nakamura and Jon Steinsson for helpful discussions and comments. I would like to thank the Social Sciences and Humanities Research Council of Canada (SSHRC) for financial support. Contact information: Department of Economics, Harvard University, Littauer Center, Cambridge MA 02138. E-mail: nakamura@fas.harvard.edu. Homepage: http://www.harvard.edu/˜ nakamura.

1 Introduction

In a report on real GDP measurement by the US Bureau of Economic Analysis (BEA), Robert Parker and Jack Triplett write,

> "It has long been recognized that fixed-weight measures misstate economic growth as one moves further from the base period. As a result, real GDP growth was understated for periods before the base period and overstated for periods after the base period."(Parker and Triplett, 1997)

Indeed, the dominant explanation for divergences among measures of real GDP in the index number theory literature is the concept of consumer substitution bias. Intuitively, substitution bias arises when consumers with fixed preferences "substitute" away from goods whose prices are rising. The well-known implication of consumer substitution bias is that the the difference between the fixed-weighted Laspeyres and chain-weighted Fisher indexes should be negative before the base period and positive after the base period. (See Section 2 for an explanation of the logic of consumer substitution bias.)

Figure 1 graphs these differences for five-year moving average growth rates, for a number of different base periods.[1] Figure 3 presents an empirical puzzle from the perspective of the quote presented above. The consumer substitution bias story does not work well for vintages of data aside from the fixed-1987 weighted series, which conforms reasonably well to the consumer substitution bias pattern. Indeed, this is the series that has often been presented in the media confirming the consumer substitution bias story. To take a particular example, the fixed-1958 weighted series is almost always negative, except for the early 1950's– when it is supposed to be negative according to the consumer substitution bias story. Indeed, *all* of the series exhibit large positive divergences in the early 1950's.

This paper both documents this empirical puzzle and proposes an explanation based on the role of demand shocks in provoking fluctuations in relative prices. The literature on real GDP measurement has traditionally taken price changes as given– framing the analysis in terms of how a consumer responds to exogenous changes in prices. This approach skirts the issue of why the price fluctuations arose in the first place. Implicitly, the assumption is that price changes arise from factors beyond the consumer (and thus outside of the

[1] I compare the average growth rates because they provide more insight into the issue of average differences in growth rates than the highly volatile raw data. The data used to construct this graph– and the caveats thereof– are discussed in Section 3.

model), such as technological innovation. In jargon, real GDP measurement proceeds under the assumption that relative price changes result from supply shocks rather than demand shocks.

Allowing for demand shocks in a model of the economy has important implications for real GDP measurement. First, one of the best known lessons from the index number theory literature is the idea that consumer substitution bias leads to an upward bias in the fixed-weighted Laspeyres measure of real GDP. Consumer substitution bias was presented as a primary motive for the BEA's decision to switch from the fixed-weighted Laspeyres to the chain-weighted Fisher measure of real GDP in 1996. Demand shocks turn this theoretical relationship on its head as I discuss in Section 4. Furthermore, from an empirical perspective, consumer substitution bias is by no means ubiquitous in national accounts statistics, as I discuss in Section 3.

This section of the paper also adds to the theoretical literature by investigating the conditions under which standard index number theory results can be applied to the case of aggregate investment. This is an important issue since, as I discuss below, a large part of the difference between alternative measures of real GDP arises from investment goods.

Second, the 1996 Boskin Commission reforms not only switched from the Laspeyres index to the Fisher index as the featured measure of real GDP, but also switched from a "fixed-base" index to a "chain-weighted" index. The latter type of index "chains" together growth rates for adjacent years based on a moving base period, in a sense that I make precise later in the paper. A comparison between fixed-weighted and chain-weighted indexes for US real GDP shows that the chain-weighted Fisher growth rate often lies significantly above the fixed-weighted version of the Fisher index. Moreover, the extent of the "bias" in the chain-weighted measure has changed considerably over the past 50 years.

It has long been known in the index number theory literature that chain-weighted indexes could "drift" away from their fixed-weighted counterparts.[2] Although general theorems on this issue are limited, numerical simulations of chain-bias generally suggest that the problem is small.[3] As in the case of consumer substitution bias, however, this analysis has been carried out under the implicit assumption of fixed, homothetic preferences. Yet, chain-bias can be very large when demand shocks are present, as I show in numerical simulations in Section 6.

[2] For example, Szulc (1983) or Forsyth and Fowler (1981) are classic references from this literature.

[3] Balk (1989) presents a number of useful results relating to chain bias.

Third, demand shocks present a major obstacle to interpreting "superlative" measures of real GDP, such as the chain-weighted Fisher index, as measures of the "utility value" of output– or more generally, as measures of theoretical output aggregators. "Superlative" indexes such as the Fisher index are output indexes, that are shown to correspond exactly to changes in utility for a particular class of utility functions. However, these results are derived under the assumption of fixed preferences, and do not hold in general if demand shocks are present.

Nevertheless, there is an intimate connection between the utility-value concept and modern quantity indexes used to measure real GDP. This connection arises from the use of prices as weights in real GDP statistics. As a consequence, it is possible to ascribe a utility-based interpretation to real GDP statistics even in the presence of demand shocks. I present this interpretation in Section 6. Although the results in both Sections 5 and 6 are derived for the case of households as the final demanders of consumption goods, the results are also directly applicable to the case of firms as the final demanders of investment goods, as I discuss in Section 4.

This empirical portion of this study is closely related to the small literature examining the impact of consumer substitution bias on the US Consumer Price Index. Some important papers in this literature are Aizcorbe and Jackman (1993), Braithwait (1980), Manser and Mcdonald (1988) and Shapiro and Wilcox (1996a). These papers use data on prices and quantities for consumption goods, at varying levels of aggregation, to assess the extent of consumer substitution bias in consumption data. The present paper uses a rather different empirical approach from the existing literature, making use instead of published real GDP statistics of different vintages. (See Section 3 for the details of the empirical approach).

The advantage of the empirical technique used in this paper is that it allows for a long-term analysis of the substitution bias phenomenon for all of the major components of real GDP. None of the papers in the literature described above analyze non-consumption data, and most cover a shorter time period than the present analysis. These are potentially important advantages since the investment and government expenditure components of real GDP have been major sources of divergence between alternative measures of real GDP. Moreover, the empirical results presented here suggest that the nature of this "bias" has changed considerably over time, suggesting that it is problematic to draw general inferences about consumer substitution bias from analyses of the recent data alone.

The disadvantage of the empirical technique used in this paper is that the comparison of alternative vintages of data incorporates the effects of some definitional changes in real GDP as well as changes in the index number formulae. I discuss this issue in more detail in Section 3

The paper proceeds as follows. Section 2 provides an overview of common empirical measures of real GDP. Section 3 presents the empirical evidence on divergences among alternative measures of real GDP statistics. Section 4 shows that demand shocks lead to the opposite sign of divergence between Laspeyres and Fisher indexes for real GDP than supply shocks given a linearly homogenous output aggregator. Section 5 discusses the implications of demand shocks for the magnitude of "chain bias" in real GDP statistics. Section 6 shows how the theorems associated with superlative indexes break down in the presence of demand shocks, and presents a utility-based interpretation of aggregate consumption that is robust to these factors. Section 7 concludes.

2 Empirical Measures of Output

Most national statistical agencies use some variation on the Laspeyres, Paasche or Fisher quantity index to compute real GDP.[4] The quantity indexes are calculated from data on the prices and quantities of disaggregated goods. The indexes are usually defined in terms of a base and current year. The Laspeyres and Paasche indexes are ratios of weighted averages of quantities in the current and base periods. The Laspeyres index, Q_L, uses the base year prices as weights,

$$Q_L = \frac{\sum_{i=1}^{M} p_i^0 y_i^1}{\sum_{i=1}^{M} p_i^0 y_i^0}, \tag{1}$$

where p_i^t denotes the price of good i in period t, y_i^t denotes the quantity of good i in period t, the base period is time 0 and the current period is time 1. The Paasche index, Q_P, uses the current year prices as weights,

$$Q_P = \frac{\sum_{i=1}^{M} p_i^1 y_i^1}{\sum_{i=1}^{M} p_i^1 y_i^0}. \tag{2}$$

[4] A little algebra shows that dividing nominal GDP by a Laspeyres price index gives a Paasche quantity index (defined below); dividing nominal GDP by a Paasche price index gives a Laspeyres quantity index; dividing nominal GDP by a Fisher price index gives a Fisher quantity index. See Diewert and Nakamura (2002b) for a derivation of these results. To a macroeconomist, it may seem a little strange to define the quantity index directly, rather than deflating nominal GDP by a price index to compute real GDP, but since we are interested in the properties of real GDP, it is simpler to deal with the formulae for the quantity indexes directly, rather than viewing real GDP as deflated nominal GDP.

The Fisher index Q_F is the geometric mean of the Laspeyres and Paasche indicies,

$$Q_F = (Q_L Q_P)^{0.5}. \tag{3}$$

Real GDP for the current year is calculated by multiplying the quantity index by the dollar value of output in the base year. (The quantity indexes themselves are unit-free ratios.)

There are two main approaches to generalizing the two-period quantity indexes to produce longer time series. The first approach is known as the the "fixed-weighted" approach. The fixed-weighted Laspeyres index is found by calculating the two-period Laspeyres index for every period, always using the same designated "base period" for period "0". In words, the fixed-weighted Laspeyres index is the expenditure that would be required, in a given period, to purchase the entire output of the economy at base period prices. The BEA used the fixed-weighted Laspeyres index to compute real GDP statistics until 1996.[5] The price-weights for the fixed-weighted series were updated at periodic intervals, in a procedure known as "rebasing".

The second approach is known as "chain-weighting." Chain-weighted indexes are constructed by computing the two-period index for every pair of *adjacent* years, and then multiplying the two-period indexes together to compute the chain-weighted index for the entire time series. For example, the chain-weighted Fisher index for period 3 is defined as the two-period index for periods 1 and 2, multiplied by the two-period index for periods 2 and 3.

In 1996, the BEA switched from a fixed-weighted Laspeyres to a chain-weighted Fisher index as its featured measure of real GDP. One of the main arguments for the switch to chain-weighting was that it eliminated the need for rebasing– providing a unified picture of macroeconomic history (rather than a picture revised at each rebasing). It is important to note, however, that the BEA's chain-weighted Fisher indexes of real GDP and its components are chained on an *annual* basis, though the statistics are released at a quarterly frequency. In particular, the growth rates for the first two quarters of a given year are calculated using price weights from the previous and current years as p_0 and p_1 in the Fisher index formula, while the growth rates for the last two quarters of the year are calculated

[5] The U.S. is actually somewhat unusual in having traditionally used the fixed-weighted Laspeyres index to compute national accounts statistics. A form of Laspeyres chain-weighting, in which fixed-weighted Laspeyres indexes of different vintages are spliced together at periodic intervals, has traditionally been used to compute real GDP in many countries outside the US. See Young (1992) for a discussion of international procedures to compute real GDP.

using price weights from the current an subsequent years.[6]

In the multi-period case, the convention in the literature is to refer to *all* fixed-weighted indexes as fixed-weighted Laspeyres indexes. A confusing feature of this terminology is that, in general, the base period for fixed-weighted Laspeyres quantity indexes produced by national statistical agencies lies somewhere in the middle of the time series. For example, the fixed-1972 weighted quantity index (a fixed-weighted Laspeyres index with base period 1972) runs from 1948-1984.

I follow the existing empirical index number theory literature by defining a fixed-weighted Laspeyres index with base period $t*$ as,

$$Q_{t*} = \frac{\sum_{i=1}^{M} p_i^{t*} y_i^1}{\sum_{i=1}^{M} p_i^{t*} y_i^0}, \tag{4}$$

However, the reader should be aware that the theoretical results in Section 4 follow the naming convention in the theoretical literature based on the two-period indexes Q_L, Q_P and Q_F defined above.

The theoretical results in Section 4 can be translated into the language of multi-period indexes by recognizing that the fixed-weighted indexe Q_{t*} is a Paasche index *before* the base period $t*$ and a Laspeyres index *after* the base period. For example, the fixed-1972 weighted index is a *Paasche* index before 1972 and a *Laspeyres* index after 1972.

3 Empirical Evidence

Even a rather crude look at the data shows that the choice of how to measure real GDP matters. Table 1 presents 10-year averages of labor productivity growth for the fixed-1987 weighted and chain-weighted Fisher series.[7] The differences between the numbers for the different measurement approaches turn out to be quite substantial. For example, the fixed-1987 weighted series shows a considerably sharper reacceleration of growth during the early 1990's than the chain-weighted Fisher series. According to the fixed-weighted Laspeyres measure, average annual labor productivity growth dropped from 1.80 in the 1970's to 0.69 in the 1980's, and then reaccelerated to 1.73 in 1990-1995. According to the chain-weighted

[6] The formula for real GDP statistics is actually even slightly more complicated than this, since the quarterly growth rates are adjusted to precisely equal the annual growth rate calculated using a chain-weighted annual index. See BEA (1998) for a discussion of this procedure. This is essentially an adjustment for chain bias.

[7] The statistics I discuss here are for the traditional measure of labor productivity, real GDP divided by hours of work.

Fisher measure, however, average labor productivity growth dropped from 2.24 in the 1970's to 1.02 in the 1980's, and then reaccelerated to only 1.24 in 1990-1995.[8]

The choice of measurement approach also affects other stylized features of the data in important ways. In particular, the switch to Fisher chain-weighting made the real GDP series considerably more volatile relative to the fixed-1992 weighted series. Parker and Triplett (1997) note that "for the five economic expansions between 1960 and 1990, the fixed (1992) weight measure of real GDP understates growth by an average annual rate of 0.8 percentage point per quarter" while "using the fixed (1992) weight measure, the annual rate of decline in real GDP during the six contractions between 1960 and 1991 is understated by an average of 0.6 percentage point."[9] This is another piece of empirical evidence that does not fit with the consumer substitution bias story, according to which the Fisher index should grow more quickly relative to the fixed-1992 weighted index during *both* booms and busts, instead of more quickly during booms and more slowly during busts as described above.

Moreover, these average differences mask some much larger differences in growth rates for individual years. The largest difference between the different measures of real GDP probably occurred in the aftermath of WWII, when the fixed-1987 weighted real GDP measure dropped by 25% between 1944 and 1947, while the chain-weighted Fisher series dropped by only 13%– a difference of 12 percentage points![10]

The dominant explanation for divergences among measures of real GDP in the index number theory literature is the concept of consumer substitution bias. Intuitively, consumer substitution bias arises because the consumer tends to substitute away from buying products whose prices increase– causing price increases to be associated with disproportionate decreases in quantity. Consumer substitution bias is generally interpreted as implying empirically that the difference between the fixed-weighted Laspeyres and chain-weighted Fisher indexes should be negative before the base period and positive after the base period. (See, for example, the Parker and Triplett (1997) quote in the introduction.)

The logic for this relationship is simple. Consider a simple example with two periods

[8]These productivity statistics do not incorporate the effects of the 1999 Comprehensive Revision but rather reflect primarily the effects of the switch to Fisher chain-weighting. See the discussion at the end of this Section.

[9]Indeed, the Fisher index appears to be more volatile than all of the fixed-weighted real GDP series.

[10]These statistics are from Landefeld and Parker (1997). Landefeld and Parker (1997), note that the divergence reflects "the use of 1987 prices for defense equipment rather than the low postwar prices for defense equipment". They go on to note that the decline in prices was associated with "the post-World War II demobilization and the associated sharp cutbacks in defense spending"– a massive demand shock.

and two goods. The Laspeyres index for period 1 growth is defined,

$$Q_{t*} = \frac{p_1^{t*} y_1^1 + p_2^{t*} y_2^1}{p_1^{t*} y_1^0 + p_2^{t*} y_2^0}, \qquad (5)$$

where the base period $t*$ is either 0 or 1. The Fisher index for this example is a geometric average of the two fixed-weighted indexes,

$$Q_F = (Q_0 Q_1)^{0.5}. \qquad (6)$$

Suppose the price of good 1 increases from period 0 to 1. Therefore, Q_1 places a higher weight on good 1 growth than Q_0, which uses the original price weights. According to the substitution bias logic, the increase in the price of good 1 causes consumers to substitute away from good 1, causing relatively slow consumption growth for good 1.

Thus, the higher weight on good 1 in Q_1 leads to slower aggregate growth rate for Q_1 than Q_0. The Fisher index growth rate lies between the growth rates for Q_1 and Q_0 since it is defined as a geometric average of the two indexes. As a consequence, the Fisher index grows less quickly than Q_0 but more quickly than Q_1. More generally, the fixed-weighted index grows more quickly the Fisher index when the base period is in the past (i.e. for Q_0) and more slowly than the Fisher index when the base period is in the future (i.e. for Q_1. In other words, we have the empirical implication described at the beginning of this Section: the difference between the fixed-weighted Laspeyres and chain-weighted Fisher indexes should be negative before the base period and positive after the base period.

The ideal way of investigating the empirical validity of this idea would be to compare different empirical quantity indexes applied to the same quantity and price data. It is not, however, possible to directly obtain data on alternative quantity indexes from publically available data sources. One can come close to comparing alternative empirical quantity indexes by comparing different "vintages" of macroeconomic time series released by the BEA. In particular, this paper uses a "real-time" dataset collected by Dean Crushore and Tom Stark (Crushore and Stark, 1999) for the period 1947-2001. These data have been used considerably in the macroeconomics literature, for example by Orphanides (2001) in an analysis of "real-time monetary policy", but have not previously been used in the index number theory literature. The series I consider are the fixed-weighted Laspeyres indexes, with base periods in 1958, 1972, 1982 and 1987 and the 1998 chain-weighted Fisher index. The different measures of real GDP statistics are available in the real-time dataset because the BEA has periodically changed its measurement approach– and thus its published measure of real GDP statistics.[11] For each base period, I make use of the real-time real GDP measure from the August of the year before a fixed-weighted series was rebased.

The BEA has used a very similar data set in its own analysis of the effects of different index number formulae on real GDP measures, though the BEA dataset covers a shorter period.[12]

The main empirical puzzle that this paper sets out to explain is presented in Figure 1. Figure 1 graphs the difference between the five-year moving average real GDP growth rates for the fixed-weighted Laspeyres and chain-weighted Fisher indexes.[13] According to the consumer substitution bias story, the difference between the fixed-weighted Laspeyres growth rates and the chain-weighted Fisher growth rates should be positive before the base period and negative after the base period for each series. As I note in the introduction, the consumer substitution bias story does not work well for vintages of data before the most recent fixed-1987 weighted measure. For example, according to the standard consumer substitution bias story, the difference for the fixed-1958 weighted series should be negative before 1958 and then become positive in the early 1960's. Instead, the difference for the fixed-1987 weighted series is positive in the early 1950's, and then almost always negative thereafter.

The standard consumer substitution bias story also does not provide a good explanation of the differences among alternative measures of real GDP components. Figures 2-4 present the same chart as Figure 1 for aggregate investment, government expenditure and consumption respectively. None of the figures are particularly consistent with the consumer substitution bias pattern overall. In the case of aggregate investment, the differences among alternative measures of real GDP statistics are quite large– the difference in average growth rates grows to over 3 percentage points for the most recent fixed-weighted series. However, only the most recent series have roughly the pattern associated with consumer substitution bias for aggregate investment. Similarly, in the case of real government expenditure, the

[11] In particular, the fixed-1958 weighted series was used from 1958 to January 1976, the fixed-1972 weighted series was used from January 1976 to December 1985, the fixed-1982 weighted series was used from December 1985 to November 1991, the fixed-1987 weighted series was used from November 1991 to January 1996, and the chain-weighted Fisher series has been used since then.

[12] I thank Marshall Reinsdorf for providing me with an unpublished BEA dataset that has been used to evaluate the effects of index number formula changes. The BEA data cover a smaller sample of years amd vintages than the real-time dataset. The results for these years, however, are very similar to the results for the real-time dataset used in this paper.

[13] Again, I compare the average growth rates because they provide more insight into the issue of average differences in growth rates than the highly volatile raw data.

differences among the growth in alternative quantity indexes are also quite large– on the order of 2 to 4 percentage points. However, the fixed-weighted Laspeyres measures grow considerably more rapidly than the chain-weighted measure during the early 1950's– the opposite of the usual substitution bias story. Finally, in the case of aggregate consumption, the size of the differences are relatively small for aggregate consumption (note the different scale of the graph for this case). However, once again, the standard substitution bias story does not hold.

In the following sections, I present some explanations of these anomalies based on the role of demand shocks and non-homothetic preferences in generating economic fluctuations. I also provide some implications of demand shocks for the interpretation of real GDP statistics.

Let us first, however, discuss some caveats associated with using the real-time data to evaluate the effects of index number formula changes. The main concern associated with using comparisons among data vintages as a proxy for comparisons among different index number formulae is that the comparison among vintages also incorporates the effects of various other types of revisions to real GDP data.

There are three main types of revisions to real GDP data once the data have been released in "final" form (as in the real-time dataset). First, there are revisions associated with changes in the index number formulae. In particular, These revisions take place during "Comprehensive Revisions" to the National Income and Product Account (NIPA) data, which have occurred approximately every 5 years since the 1980's.[14] Second, there are revisions associated with new data sources becoming available. These revisions are implemented both during yearly "Annual Revisions", which mainly update data sources revise data from the past few years of a series, and during Comprehensive Revisions, which may incorporate data for as much as the last 10 years of a given data series. Third, there are changes in the definition of the BEA's featured measure of output, such as the switch from GNP to GDP as the featured measure of output in the 1991 Comprehensive Revision. The treatment of new goods and quality-adjustment issues has also changed somewhat over time since the BEA began using hedonic methods for a very limited portion of real GDP in 1987. More recently, the 1999 Comprehensive revision introduced major changes in the BEA's approach to measuring new goods and quality improvement.

Although the latter two types of revisions to real GDP are certainly a concern in using

[14] Specifically, Comprehensive Revisions occurred in 1976, 1985, 1991, 1996, 1999 and 2004.

different vintages of real GDP data to examine the effect of changes in index number formulae, it is generally acknowledged that the larged source of revisions to the growth rate of real GDP statistics. For example, the BEA's report on the 1991 Comprehensive revision comments that "the downward revisions in the growth rates of real GDP...are mainly accounted for by the shift in the base period." (BEA, 1991) On the same note, Crushore and Stark (2001) comment that revisions to real GDP growth tend to be larger than revisions to nominal GDP growth since only the former type of statistics incorporates the effects of index number formuala changes.

More specifically, the effects of pure data revisions are typically concentrated near the end of a given series, and thus should not affect comparisons for earlier periods. On the issue of definitional changes to real GDP, I make use of the 1998 version of the chain-weighted real GDP measure to avoid incorporating the revisions to real GDP associated with an increased use of hedonic methods during the 1999 Comprehensive Revision. The hedonic methods used before 1996 were limited in scope and were not applied retroactively, and thus should not affect years before 1987.

It should also be noted that there are significant advantages to the empirical approach in this paper relative to previous empirical work on substitution bias that used data on prices and quantities for various categories of consumption goods. In particular, previous studies of substitution bias have been limited to to studying substitution bias at an aggregate level, since disaggregated data on prices and quantities are not available. In addition, previous studies have focused only on the consumption component of real GDP.

Nevertheless, contemporary statistics on alternative measures of real GDP statistics would clearly be extremely helpful in carrying out this type of analysis. Furthermore, the potential impact of definitional or other non-formula based revisions to real GDP on growth rates is an important consideration in assessing the empirical results in this paper, particularly in the case of very small differences in growth rates.

4 Effect 1: Demand Shocks and Substitution Bias

Consumer substitution bias results from the way a utility-maximizing consumer responds to a price change. Namely, the consumer tends to substitute away from buying products whose prices increase– causing price increases to be associated with disproportionate decreases in quantity. The same type of intuition can be applied to the case of a firm facing changes in factor prices– motivating the term "final demander substitution bias" in the case of the aggregate investment component of real GDP. The Laspeyres output index, which incorporates prices from before the price increase, places a higher price weight on products that subsequently have high growth than the Fisher output index, which incorporates prices from *after* the price change. (See Section 3 for the logic of this relationship.) The Fisher index grows less quickly since it tends to weight rapidly growing items by lower prices. The same type of argument can be applied to the behavior of profit-maximizing firms facing exogenous changes in input prices.

Borkiewitz (1924) formalizes this notion, showing that,

$$\frac{Q_P - Q_L}{Q_L} = \operatorname{cov}(\frac{p_1}{p_0}, \frac{q_1}{q_0}), \tag{7}$$

where cov denotes the covariance between the price and quantity ratios. Thus, the sign of the difference between the Laspeyres and Fisher measures of output growth depends on the covariance between price and quantity changes. The standard consumer substitution bias result is that since consumers substitute away from goods with rising prices, $\operatorname{cov}(p, q)$ is negative. Therefore, $Q_F < Q_L$. The result that $Q_F < Q_L$ can also be derived in a utility-maximizing framework, given fixed homothetic preferences. (See below.)

Although consumer substitution bias is often interpreted as saying that the chain-weighted index grows more quickly before the base period and more slowly after the base period, Borkiewitz's result (7) really only applies to comparisons between two periods where the first period is the base year of the fixed-weighted index. Thus, a useful complement to Figure 1 is to compare average growth rates for alternative measures of real GDP statistics in the period immediately following the base years for the fixed-weighted indexes.[15]

In particular, Table 2 compares the growth in the fixed-weighted Laspyeres indexes to the growth in the chain-weighted Fisher index in the 5 years following the base years 1958, 1972, 1982, and 1987. Table 2 presents this information for the aggregate investment, consumption and government expenditure series. According to the consumer substitution bias story, the Fisher index should always grow more quickly than the Laspeyres index over these periods. The table shows, however, that the relationship is often reversed. Moreover, the differences are numerically large in the case of both the aggregate investment and

[15] The theory of consumer substitution bias does not actually imply any prediction for the relative growth rates from, say, 1977 and 1982 for the fixed-1972 weighted versus the chain-weighted Fisher measures. The problem is that the fixed-1972 weighted measure reflects prices in 1972, and thus an increase in prices relative to 1977 levels may in fact be a decrease in prices relative to 1972 levels.

government expenditure growth rates.

In Section 5 I suggest that there may be significant differences between the chain-weighted Fisher index and the direct Fisher index defined in Section 2. In order to deal with this issue, Tables 3 and 4 present evidence on the divergence between alternative quantity indexes that does not depend on the chaining procedure for the Fisher index.

Table 3 presents the difference fixed-weighted Laspyeres indexes to the growth in the chain-weighted Fisher index in the year before and after the base years 1958, 1972, 1982, and 1987.[16] The chain-weighted Fisher index is essentially identical to the direct Fisher index over these one year periods. According to the substitution bias story, the difference between the fixed-weighted Laspeyres and chain-weighted Fisher index should be negative before the base period and positive after the base period. The table reiterates the previous result that, at least in the short term, this pattern often does not hold. For example, in the case of government expenditure, the divergence is large and negative following the 1958 base year when it should be positive according to the consumer substitution bias story.

Table 4 presents a comparison *among* fixed-weighted indexes with differnt base periods. Over the interval between base periods, the fixed-weighted indexes of different base periods can be interpreted as basic Laspeyres and Paasche indexes. Recall that according to the consumer substitution bias story, the Laspeyres index grows more quickly than the Paasche index. Thus, the "Laspeyres" index in Table 4 is the fixed-weighted index with the base period at the beginning of the time interval over which growth is measured; while the "Paasche" index is the fixed-weighted index with the base period at the end of the time interval. Table 4 shows that over the longer time periods considered in the table, the substitution bias pattern is dominant. The only cases for which the Laspeyres index grew significantly more quickly than the Paasche index are for aggregate investment over the period 1958q1-1971q4 and government expenditure over the period 1982q1-1987q4. Table 4 shows that the Laspeyres index also grews slightly more quickly than the Paasche index for imports over the period 1958-1972. However, the extent to which the Laspeyres and Paasche index evidently varies greatly over time– the difference is particularly large for the investment aggregate during the 1982-1987 period. These results suggest tentatively that the consumer substitution bias story is more empirically successful over longer time periods.

[16] Specifically, I look at the year that ends in the second quarter of the base year, and the year that begins in the third quarter of the base year. As I discuss in Section 2 the BEA's measure of real GDP is calculated using prices from the current and previous years in Q1 and Q2 of a given year; while it is calculated using prices from the current and subsequent years in Q3 and Q4 of a given year.

Let us now consider how these empirical results can be interpreted from the perspective of economic theory. The standard results on consumer substitution bias in the index number theory literature take price changes as exogenous from the perspective of the consumer.[17] In this section, I investigate how these results change if we take the equilibrium perspective that price changes *result* from demand and supply shocks. First consider the implications of demand shocks for the correlation between prices and quantities. In the Economics 101 diagram of downward sloping demand and upward sloping supply, a shift in the supply curve with a fixed demand curve causes an increase in quantity and a decrease in price. In contrast, a shift in the demand curve causes increases in both quantity and price. Thus, the Fisher index may grow *either* more quickly or more slowly than the Laspeyres index depending on whether the economic fluctuations are caused by demand or supply shocks. I present a simple worked out general equilibrium model in Appendix A with the property that demand shocks lead to price and quantity increases, while supply shocks lead to price increases and quantity declines.

An anecdotal example of a demand shock that plays a large role in generating Figure 1 is the military build-up associated with the Korean war. The military build-up during the early 1950's caused an increase in prices for defense goods (at least between the early 1950's and 1958). One can see this from the rapid growth in the fixed-1958 weighted series relative to the chain-weighted Fisher series in the early 1950's for both real GDP and government expenditure (Figure 1 and Figure 3). This behavior is hard to reconcile with fixed preferences, and fits with the accepted historical view that the military build-up during this period was essentially demand driven.

In the case of the aggregate investment series, it is more appropriate to refer to the substitution bias phenomenon as "final demander substitution bias". Clearly, the results above for the implications of alternative correlations between price and quantity changes apply whether the final demander is a consumer purchasing consumption goods or a firm purchasing investment goods. Moreover, I show below that indeed, an analogy can be drawn between a utility-maximizing household's response to consumer product price changes and a profit-maximizing firm's response to factor price changes. Thus, the results on consumer substitution bias from the "economic approach" can also be applied to the more general concept of "final demander substitution bias".

[17] For example, see Diewert's (1992) survey on index numbers in the *Palgrave* dictionary for a summary of these results.

Although it seems quite reasonable that demand and supply shocks can produce opposite correlations between price and quantity changes described above, it is useful to think about how these results can be derived from optimizing behavior of firms and households. The following propositions develop in detail the relationship between the assumption of fixed, homothetic preferences and the differences among alternative quantity indexes.[18] The first part of the proposition (1a) is a restatement of standard results in the index number theory literature on substitution bias. These results are reproduced, for example in Diewert (1992). The second part of the proposition is a restatement of results in Diewert (1973), who builds on the results of by Afriat (1967). The remainder of the results in this section have not, to my knowledge, previously appeared in the index number theory.

Proposition 1: When the consumer is the final demander (as in the aggregate consumption data), (1a) $Q_P < Q_F < Q_L$ is consistent with household optimization given stable, homothetic preferences while (1b) the opposite relationship is not.

Proof: The proofs are in Appendix B.

This proposition derives the consumer substitution bias result for the case of homothetic preferences, but also derives a converse result that shows that the sign of the difference between Q_F and Q_L tells us something about the nature of the shocks that caused the price changes. The key implication of this result is that *if we observe the empirical "anomaly" that $Q_L < Q_F$ we can conclude either that the demand function shifted or that preferences are non-homothetic.* Although the homotheticity assumption is clearly an oversimplification, it is almost always used in the theoretical literature in which aggregation appears because it provides a simple justification for aggregation of prices and quantities. For example, in the macroeconomics literature, the Dixit-Stiglitz consumption index is a well-known consumption aggregator with this property.

We can now prove an analagous result for the case of the aggregate investment series, where firms are the final demanders of investment goods. The simplest way of allowing for aggregation among capital inputs is to assume that the firm's production function justifies a linearly homogenous investment aggregate. That is, the firm's production function, $G(L_t, K_t)$ depends on a capital aggregate K_t, and we make the strong assumption that $K_{t+1} = K_t(1-\delta) + I_t$, where I_t is a linearly homogenous investment aggregate and δ is

[18] I thank W. Erwin Diewert for extensive and detailed comments on this section of the paper.

the depreciation rate– that is, the capital stock is separable in the aggregate investment for each period. In this case, the following proposition holds. It is important to note that these additional assumptions on the form of the capital aggregate are necessary to transfer the standard results on consumer substitution bias to the case of aggregate investment.

Proposition 2: When the firm is the final demander (as in the aggregate investment data), (2a) $Q_P < Q_F < Q_L$ is consistent with firm optimization given a stable, linearly homogenous investment aggregate while (2b) the opposite relationship is not.

Proof: The proofs are in Appendix B.

This proposition extends the usual "consumer substitution bias" result to the case of firms purchasing investment goods. Moreover, similar to the proposition for the *consumer* substitution bias case, this proposition shows that if we observe the empirical "anomaly" that $Q_F < Q_L$ we can conclude either that the investment aggregator has shifted *or* the linear homogeneity assumption fails to hold.

The propositions so far have illustrated that the relationship $Q_P < Q_F < Q_L$ is *not* consistent with optimizing behavior given a stable consumption aggregator on the part of households or firms and standard homotheticity assumptions. The following propositions show that this "anomaly" can, however, be rationalized by optimizing behavior on the producer side of the economy, given shifts in demand. In particular, let us assume that the production possibilities frontier has the form,

$$S = (y, x) : g(y) = f(x), \tag{8}$$

where y is a vector of outputs and x is a vector of inputs, and $g(y)$ is a linearly homogeneous function. In this case, the firm's output has the form,

$$q_t^i = g(y)h(p_t), \tag{9}$$

where q_t^i is output of good i in period t and h is a function of prices p. That is, for a given level of prices, the firm's outputs simply scale up with the level of $g(y)$.[19] Thus, $g(y)$ is a natural output aggregator.

[19]This a direct analogy to the case of homothetic preferences. The firm can always achieve a given level of output $g(y)$ by scaling up the outputs for the case where $g(y) = 1$. Moreover, the resulting scaled up outputs must be the revenue maximizing outputs for a given level of output $g(y)$ since revenue is linear in y.

The following propositions show that the relationship $Q_P < Q_F < Q_L$ *is* consistent with optimizing behavior given a stable output aggregator $g(y)$– while the opposite relationship is not.

Proposition 3: When the consumer is the final demander, (3a) $Q_L < Q_F < Q_P$ is consistent with producer optimization given stable production technology and a linearly homogenous output aggregator while (3b) the opposite relationship is not.

Proof: The proofs are in Appendix B.

Notice that this proposition makes no assumptions about the *demand* side of the economy, unlike Propositions 1 and 2 which assumed homothetic, stable preferences. The proposition assumes implicitly that any fluctuations in relative price arise from *supply* shocks or non-homothetic aggregators on the supply side of the economy. The key implication is that *if we observe the relationship $Q_L < Q_F$ we can conclude either that the supply function shifted or that the output aggregator is non-homothetic.* The approach used to prove this result is analagous to the approach taken in the literature on theoretical "output price indexes" as discussed, for example, in Alterman, Diewert and Feenstra (1999).

Once again, it is possible to derive an analogous result for the case where the outputs of one firm are inputs into another firm's production process– i.e. in the case of aggregate investment. In this case, however, the demand shocks refer to changes in the demand of other firms for the investment goods.

Proposition 4: When the firm is the final demander, (4a) $Q_L < Q_F < Q_P$ is consistent with producer optimization given stable technology of the capital-producing firm and a linearly homogenous output aggregator for the capital-producing firm while (4b) the opposite relationship is not.

Proof: The proofs are in Appendix B.

The main conclusion of these results is to show that if economic fluctuations are caused by both demand and supply shocks, or if preferences or the capital aggregator are non-homothetic, then it is quite possible to generate either $Q_L < Q_F < Q_P$ or $Q_L > Q_F > Q_P$. Furthermore, if we are willing to make the homotheticity assumption, then the sign of the "bias" in the fixed-weighted Laspeyres index can be used to infer whether the economic fluctuations were generated by demand or supply shocks.

There are a number of potential reasons why demand shocks may occur (or appear to occur) aside from "true" changes in preferences. One likely reason is the presence of durable goods and habit formation. In the case of consumers, the presence of durable goods or habit formation implies that the utility function is not fully separable in the consumption good across time periods. In the case of firms, the assumption that the capital aggregate is separable across different periods in time is clearly an oversimplification given that many capital goods are highly durable. Both of these factors may be important in explaining the apparent "demand shocks" discussed in this section. Another potential source of demand shocks are shifts in the composition of government expenditure, such as the military expenditures associated with wars discussed above.

5 Effect 2: Demand Shocks and the Chain Drift Problem

The index number theory literature has historically focused on studying the relationships among the basic empirical quantity indexes– the Laspeyres, Paasche and Fisher quantity indexes, which I define in Section 2. However, as I note in the introduction, real GDP is not actually calculated using any of these formulae in the US. Rather, the US BEA has used the chain-weighted Fisher index as the featured measure of real GDP statistics since 1996. It is assumed in much of the literature, such as the literature on "consumer substitution bias" cited in the introduction, that the chain-weighted Fisher index closely approximates the direct Fisher index defined in Section 2 by equation (34).

This viewpoint has some theoretical support. Although it is well-known in the index number theory literature that there is the potential for chain-weighted indexes to "drift" away from their direct counterparts, there is relatively little theory on the nature of chain drift. Some important articles on the topic of "chain-drift" are Szulc (1983) and Forsythe and Fowler (1981). These articles present a number of insightful examples of plausible examples of price and quantity paths and calculates the behavior of Q_F and Q_F^{ch} for these examples. The main lessons from these examples are that chain-drift is likely to be a problem if prices and quantities "bounce" up and down– and even then the effect is likely to be relatively small. In particular, in all of these examples Q_F^{ch} turns out to lie between the direct Laspeyres and Paasche indexes. A notable exception is Feenstra and Shapiro (2001) which presents an example of massive chain bias. Balk (1989) also presents a number of number of results that are relevant to the issue of chain bias, as I discuss below.

How well does the main prediction of the literature– that "drift" in the chain-weighted Fisher index relative to the direct Fisher index tends to be small– line up with the data on real GDP statistics? Although the BEA does not publish the direct Fisher quantity index, it is possible to calculate the direct Fisher index over certain time periods using the published fixed-weighted series. For example, the fixed-1972 weighted index and the fixed-1982 weighted index can be used to calculate the direct Fisher index for the period 1972-1982.

Table 3 presents a comparison of the direct Fisher and chain-weighted Fisher growth rates for the periods 1958-1972, 1972-1982, and 1982-1987. The table presents the comparison for the aggregate investment, consumption and real government expenditure components of real GDP. The table suggests that the chain-weighted Fisher index has often drifted *substantially* away from its fixed-weighted counterpart. This "drift" in the chain-weighted Fisher index contributes considerably to the anomalies in Figure 1 that I discuss in Section 3.

Can demand shocks explain this phenomenon? A major barrier to understanding chain bias has been the intractibility of the formula for chain bias, as a consequence of the complicated form of chain-weighted index numbers. In order to derive a more tractable formula, let us consider the following linear approximation. The chain-weighted Laspeyres index is defined[20],

$$Q_L^{ch} = \prod_{t=0}^{T} \frac{\sum_{i=1}^{M} p_i^t y_i^{t+1}}{\sum_{i=1}^{M} p_i^t y_i^{t+1}}. \tag{10}$$

Thus,

$$\begin{aligned} \log Q_L^{ch} &= \sum_{t=0}^{T} \log \frac{\sum_{i=1}^{M} p_i^t y_i^{t+1}}{\sum_{i=1}^{M} p_i^t y_i^t} && (11) \\ &= \sum_{t=0}^{T} \log(\sum_{i=1}^{M} s_{i,t} G_{i,t}) && (12) \\ &\approx \sum_{t=0}^{T} \sum_{i=1}^{M} s_{i,t} g_{i,t}, && (13) \\ & && (14) \end{aligned}$$

where $G_{i,t} = \frac{y_t}{y_{t-1}}$ is the growth between periods $t-1$ and t for good i, $g_{i,t} = G_{i,t} - 1$ is the

[20]I am abstracting here from the fact the BEA's chain-weighted Fisher indexes of real GDP and its components are chained on an *annual* basis, though the statistics are released at a quarterly frequency.

corresponding growth rate, and $s_{i,t} = \frac{p_{i,t}y_{i,t}}{\sum_{i=1}^{M} p_i^t y_i^t}$ is the expenditure share on good i in period t. The first line follows from a standard algebraic manipulation and the second line follows from computing a Taylor series approximation of $\log(x)$ around the point $x = 1$ where x represents the sum $\sum_{i=1}^{M} s_{i,t}G_{i,t}$.[21]

The approximation (16) is quite useful in understanding the source of chain drift. Let us further rewrite (16) as,

$$\log Q_L^{ch} \approx \sum_{t=0}^{T}\sum_{i=1}^{M} s_{i,t}g_{i,t}, \tag{15}$$

$$= T\sum_{i=1}^{M}[\bar{s}_i\bar{g}_{i,t} + \mathrm{cov}(s_{i,t}, g_{i,t})], \tag{16}$$

where $\mathrm{cov}(s_{i,t}, g_{i,t})$ denotes the *sample* covariance between expenditure shares and growth over time and $\bar{s}_i = \sum_{t=0}^{T} s_{i,t}/T$. A similar line of reasoning shows that,

$$Q_P^{ch} \approx T\sum_{i=1}^{M}[\bar{r}_i\bar{g}_{i,t} + \mathrm{cov}(r_{i,t}, g_{i,t})], \tag{17}$$

where $r_{i,t} = \frac{p_{i,t+1}y_{i,t}}{\sum_{i=1}^{M} p_i^t y_i^t}$ and and $\bar{r}_i = \sum_{t=0}^{T} r_{i,t}/T$. Notice that $r_{i,t}$ is *not* an expenditure share– it differs from $s_{i,t}$ by the factor $\frac{p_{i,t}}{p_{i,t+1}}$.

The same approximation can be applied to the direct Laspeyres index to show that,

$$\log Q_L \approx \sum_{i=1}^{M} s_{i,0}g_{i,0,T}, \tag{18}$$

where $g_{i,0,T}$ is the growth rate from time 0 to T. If we now approximate this growth rate by the sum,[22]

$$g_{i,0,T} \approx \sum_{t=0}^{T} g_{i,t}. \tag{19}$$

then we can rewrite the direct Laspeyres index as,

$$\log Q_L \approx \sum_{t=0}^{T}\sum_{i=1}^{M} s_{i,0}g_{i,t} \tag{20}$$

$$= T\sum_{i=1}^{M} s_{i,0}\bar{g}_{i,t}. \tag{21}$$

[21] This approximation deteriorates as growth rates become large in absolute value. Notice that $\sum_{i=1}^{M} s_{i,t}G_{i,t} = 1$ if $G_{i,t} = 1$ for all i.

[22] This approximation follows, from taking a linear approximation of the expression $\log\prod_{t=0}^{T} G_{i,t}$. As before, the approximation deteriorates as the growth rates become large in absolute value.

Similarly, the direct Paasche index can be approximated as,

$$\log Q_L \approx T \sum_{i=1}^{M} r_{i,T} \bar{g}_{i,t}. \tag{22}$$

We are now ready to compute our approximate formula for the chain bias in the Fisher index. Recall from Section 2 that the Fisher index is the geometric average of the corresponding Laspeyres and Paasche indexes. Thus, we have the following expression for chain bias in the the chain-weighted Fisher index,

$$\log Q_F^{ch} - \log Q_F \approx \phi + \frac{T}{2} \sum_{i=1}^{M} [\text{cov}(s_{i,t}, g_{i,t}) + \text{cov}(r_{i,t}, g_{i,t})], \tag{23}$$

where,

$$\phi = T \sum_{i=1}^{M} \left[\frac{\bar{s}_i + \bar{r}_i}{2} - \frac{s_{i,0} + r_{i,T}}{2} \right] \bar{g}_{i,t}. \tag{24}$$

Thus, the magnitude of chain bias depends on two factors: first, the ϕ term, which depends on the difference in averages of the expenditure share terms $s_{i,t}$ and $r_{i,t}$ over the entire sample versus the endpoints 0 and T only; and second, the covariance of these terms with growth growth over time. A number of the results in the literature on chain bias can be understood in the context of expression (23). First, it is clear from these expressions that if the growth rate $g_{i,t}$ is constant then the chain-bias is small, since the covariance term is zero and $(\bar{s}_i + \bar{r}_i)/2$ is close to $(s_{i,0} + r_{i,T})/2$. This lines up with Szulc's (1983) observation that chain bias is more problematic when prices and quantities "bounce" from period to period. However, the approximation (23) also shows that chain drift is likely to occur whenever there is a significant correlation between expenditures shares and growth– even when prices and quantities do not exhibit the particular pattern of bouncing discussed in Szulc (1983).

In order to understand why theoretical examples of chain bias have generally not found big effects– and why demand shocks are potentially an important consideration for this problem– let us consider the following slight modification of an example of chain-bias from the well-known article by Szulc (1983).[23]

$$\begin{aligned} p_1 &= 14, 12, 10, 12, 14 \\ p_2 &= 10, 10, 10, 10, 10 \end{aligned}$$

[23] I have interchanged prices and quantities in Szulc's example since Szulc (1983) focused on price indicies, while the present paper focuses on quantity indicies.

$$
\begin{aligned}
p_3 &= 6, 8, 10, 8, 6 \\
q_1 &= 3, 4, 5, 4, 3 \\
q_2 &= 5, 5, 5, 5, 5 \\
q_3 &= 7, 6, 5, 6, 7,
\end{aligned}
$$

where p_i and q_i are the price and quantity respectively of good i.

More generally, consider the case where outputs return to their original level by time T i.e. $y_{i,0} = y_{i,T}$ for all goods i. Thus $Q_L = Q_P = Q_F = 1$. Thus,

$$\log Q_L^{ch} - \log Q_L = \log Q_L^{ch}, \tag{25}$$

since $\log Q_L = 0$. The same is true for Q_P. As a consequence,

$$\log Q_F^{ch} - \log Q_F \approx \frac{T}{2} \sum_{i=1}^{M} [\text{cov}(s_{i,t}, g_{i,t}) + \text{cov}(r_{i,t}, g_{i,t})]. \tag{26}$$

This expression follows directly from the expression for chain bias (23) noting that we have assumed that $\bar{g}_i$ for all i. In other words, the magnitude of the chain bias depends solely on the covariance between the expenditure share terms, $s_{i,t}$ or $r_{i,t}$, and growth. However, in Szulc's example, the expenditure shares vary relatively little,

$$
\begin{aligned}
s_1 &= 0.31, 0.33, 0.33, 0.33, 0.31 \\
s_2 &= 0.37, 0.34, 0.33, 0.34, 0.37 \\
s_3 &= 0.31, 0.33, 0.33, 0.33, 0.31,
\end{aligned}
$$

where s_i is the expenditure share of good i. As a consequence, the example produces very little correlation between output growth and shares and hence very little chain bias: $Q_F^{ch} = 0.998$ is very close to $Q_F = 1$.

More generally, suppose we assume the Cobb-Doublas expenditure function,

$$C(p, u) = \prod_{i=1}^{N} (p_i^{a_i}) u, \tag{27}$$

where the a_i's are parameters. This expenditure function implies constant expenditure shares given by a_i for each good i, Given the results above, we have the following proposition.

Proposition 5: If preferences are stable, and the expenditure function has the Cobb-Douglas form then,

$$Q_F^{ch} - Q_F \approx \frac{T}{2}[\mathrm{cov}(s_i - r_{i,t}, g_{i,t})]. \tag{28}$$

Proof: The result follows directly from the approximation to the chain bias in the Fisher index, (26), noting that expenditure shares are constant over time for the Cobb-Doublas expenditure function.

The terms on the right hand side of expression (28) tend to be small if prices adjust relatively gradually (note that $s_{i,t}$ and $r_{i,t}$ differ only by $(p_{i,t+1} - p_{i,t})/(\sum_{i=1}^{M} p_{i,t}y_{i,t})$). Thus, chain bias tends to be small given a Cobb-Douglas expenditure function. Once again, an analogous result holds for the case of firms as the final demanders of factors of production, such as investment goods. These results follow immediately by assuming a Cobb-Doublas cost function for investment goods, which implies constant shares for each type of investment good.

In contrast, if the demand function shifts over time (or if we allow for a more general specification of utility), then it is quite possible to generate significant variation in shares– and thus, significant chain bias. For example, consider the following sequence of prices and quantities.

$$\begin{aligned} p1 &= 1,2,3,4,10,9,9,8,7 \\ p2 &= 1,2,3,4,5,4,3,2,1 \\ p3 &= 1,2,3,4,5,4,3,2,1 \\ q1 &= 5,4,3,2,1,2,3,4,5 \\ q2 &= 1,1,1,1,1,1,1,1,1 \\ q3 &= 1,1,1,1,1,1,1,1,1 \end{aligned}$$

In this example, the chain bias is very large: $Q_F^c h = 1.30$ while Q_F remains 1. Just as in Szulc's example, the quantity fluctuations in this example are purely temporary. Unlike in the previous example, however, since the price of good 1 trends upward over time, its expenditure share increases dramatically. This generates a correlation between relative prices and growth, since growth is positive for the first several periods and negative for the remaining perids. The positive correlation between growth and expenditure shares generates considerable upward chain bias.

Notice that in this example, while quantities return to their original levels, relative prices do not. As a consequence, it is evident that either demand has shifted or preferences are non-homothetic. Feenstra and Shapiro's (2001) example of massive chain drift in high frequency price indexes also has this characteristic. In particular, Feenstra and Shapiro (2001) present an example in which quantity demanded tends to increase in the final weeks of a sale. In other words, quantity increases in the final weeks of a sale are not associated with contemporaneous changes in prices. The authors attribute this feature of the data to increased promotion of sale items in the final weeks of a sale.

Balk (1989) presents a related result for the case of chain-weighted Fisher price indexes.[24] Although Balk's results are derived for the case of price indexes, an analagous result applies for the case of quantity indexes. Namely, if the price weights between 0 and T affect the quantity index via the log-linear function assumed by Balk, the spread between the Laspeyres and Paasche indexes for adjacent periods is small, and all quantities return to their time 0 levels by time T then the value of the chain-weighted Fisher quantity index between 0 and T approximately equals 1 *unless* there has been a demand shock during this interval.

It has often been argued that high frequency data should be avoided in using chain-weighted indexes to avoid the chain drift problem. Since the BEA's chain-weighted Fisher indexes of real GDP and its components are chained on an *annual* basis, they do not incorporate seasonal fluctuations in prices. However, even at an annual frequency, real GDP statistics– such as aggregate investment– often have highly volatile growth rates. Moreover, it is relatively common for relative prices of particular goods to trend upward or downward over time. The approximation (23) shows that these features of the data are likely to considerably exacerbate the chain drift problem.

[24] This result is also reproduced in Balk (2004).

6 Effect 3: Demand Shocks and the Interpretation of Real GDP

The "economic approach" to index number theory provides an interpretation for aggregate statistics based on their relationship of index numbers to theoretical output aggregates.[25] In this vein, theoretical macro-models often assume that the consumer's utility depends on the Dixit-Stiglitz (1977) consumption index,

$$C = \left[\int_0^1 c(j)^{\frac{\theta-1}{\theta}} dj \right]^{\frac{\theta}{\theta-1}}, \tag{29}$$

where $c(j)$ is output of good j, and θ is a parameter describing the substitutability of the consumption goods. Notice that the index described by expression (30) is a linearly homogenous output aggregator. A similar type of approach is often applied to aggregating different types of capital goods. Indeed, in many macroeconomic models, the investment aggregate is also assumed to take the form (30).

Seminal results in the literature on the economic approach to index numbers show that theoretical aggregates can be measured exactly by certain index numbers given particular assumptions about the functional forms of the aggregates and optimizing behavior on the part of firms and households. An important result in this area is that if preferences are fixed and the expenditure function has a particular "flexible functional form", then the Fisher quantity exactly equals the ratio of utility in the comparison periods.[26] More generally, index numbers with the property that they correspond to the theoretical aggregates for expenditure functions that have the "flexible functional form" property are known as superlative index numbers. The fact that the Fisher index is a superlative index number was an important theoretical argument for its adoption by the US BEA in 1996.

In this paper I have suggested, however, that one common source of price fluctuations is shifts in the preferences of the final demander for individual goods. These shifts correspond to changes in the form of the aggregator function itself. For example, we can modify the Dixit-Stiglitz index to allow for preference fluctuations as,

$$C = \left[\int_0^1 z_j^{\frac{1}{\eta}} y(j)^{\frac{\theta-1}{\theta}} dj \right]^{\frac{\theta}{\theta-1}}, \tag{30}$$

[25] The "economic approach" to index number theory is to be contrasted with the axiomatic approach. See Diewert (1988) for a discussion of the latter.

[26] Diewert (1973, 1974, 1976) and Diewert and Wales (1992,1995) define a utility function as "flexible" if it is capable of providing second-order approximation to an arbitrary twice-differentiable linearly homogeneous utility function.

where z_j is a preference shock.

Let us first observe that if we allow for these types of demand fluctuations then the standard results on superlative indexes fail to hold. The results on superlative indexes depend crucially on the idea that the consumption aggregate, for example, has a stable form. Thus, the correspondance between theoretical output aggregates and superlative index numbers is, perhaps, not as close as one might have hoped given the empirical evidence in Section 4.

How should we then interpret real GDP statistics? One answer is to look for a theoretical output aggregator on the *producer* side, rather than the consumer side of the model.[27] Unfortunately, this type of aggregate yields an analogous set of problems associated with shifts on the producer side of the economy. An alternative approach used, for example, by Feenstra and Shapiro (2001) is simply to *define* the theoretical price or quantity index in terms of observable quantities such as expenditure. The downside of this approach is that output aggregates defined in this manner do not correspond well to the aggregate output variables that appear in theoretical macro-models, or to common intuitive interpretations of real GDP statistics.

It is useful, therefore, to consider the following middle ground between finding an exact analogue to theoretical output aggregates and defining output aggregates as ratios of observable quantities. Consider the standard household optimization problem,

$$\max U^t(x) \qquad \text{s.t.} \qquad p^t y = W^t, \tag{31}$$

where U^t is the linearly homogeneous utility function for period t, p^t is the vector of period t prices, y is a vector of consumption goods, and W is the household's wealth. The optimization problem implies the familiar first order conditions,

$$MU_j^t = \lambda^t p_j^t, \tag{32}$$

where p_j^t is the price of the j'th good in period t, MU_j^t is the marginal utility of the j'th good in period t and λ^t is the marginal utility of wealth in period t. That is, prices are proportional to marginal utilities. By the definition of the Konus price index, the Konus price index is proportional to the inverse of the marginal utility of wealth.[28] Thus, the first

[27] See Alterman, Diewert and Feenstra (1999) for a discussion of this approach.

[28] By definition, the wealth required to increase utility by dU is $P^t dU * c^0(p^0)$ where P^t is the Konus price index in period T and $c^0(p^0)$ is the cost required to purchase one unit of utility in the base period. Therefore, $P^t \propto 1/\lambda^t$.

order conditions (32) can be rewritten as,

$$MU_j^t \propto \frac{p_j^t}{P^t}, \tag{33}$$

where P^t is the Konus price index in period t, and "$\propto$" means "proportional to".

Recall from Section 2 that the Laspeyres index is defined,

$$Q_L = \frac{\sum_{i=1}^M p_i^0 y_i^1}{\sum_{i=1}^M p_i^0 y_i^0}. \tag{34}$$

Substituting marginal utilities for prices according to conditions (32) we have,

$$\begin{aligned} Q_L &= \frac{\sum_{i=1}^M \lambda^0 MU_i^0 y_i^1}{\sum_{i=1}^M \lambda^0 MU_i^0 y_i^0} \\ &= \frac{\sum_{i=1}^M MU_i^0 y_i^1}{\sum_{i=1}^M MU_i^0 y_i^0}. \end{aligned} \tag{35}$$

This expression is robust both to demand shocks and non-homothetic preferences. If preferences are homothetic, the ratio of utility in period 1 to period 0 can be rewritten as,

$$\frac{U(y^1)}{U(y^0)} = \frac{\sum_{i=1}^M MU_i^1 y_i^1}{\sum_{i=1}^M MU_i^0 y_i^0}, \tag{36}$$

due to Euler's theorem for linearly homogenous functions. A comparison between equations (35) and (36) shows that, in the case of homothetic preferences, the Laspeyres index can be interpreted as the proportional increase in household utility that *would have occurred* holding marginal utilities fixed at base period levels.

A similar argument can be used to show that the Paasche index measures the proportional increase in household utility that *would have occurred* holding marginal utilities fixed at period 1 levels. Namely,

$$Q_P = \frac{\sum_{i=1}^M MU_i^1 y_i^1}{\sum_{i=1}^M MU_i^1 y_i^0}. \tag{37}$$

Figures 5 and 6 graphical representation of this result for both the demand and supply shock cases. For simplicity, I normalize base period utility to 1 in these figures so that we can rewrite the Laspeyres and Paasche indexes simply as,

$$\begin{aligned} Q_L &= \sum_{i=1}^M MU_i^0 y_i^1 \\ Q_P &= \sum_{i=1}^M MU_i^1 y_i^1. \end{aligned} \tag{38}$$

Figure 5 presents the supply shock case, where preferences are stable and the adjustment in prices is caused by a supply shock. Moreover, I assume that the supply shock is engineered so that only one of the consumption goods increases in quantity. In contrast, Figure 6 depicts the demand shock case, where there is a shift in the utility function. Notice that the opposite relationships hold between Q_L and Q_P in Figure 5 versus Figure 6. The figure shows that intuitively, these differences arise from the fact that marginal utility falls after the quantity increase in the supply shock case (due to diminishing marginal utility) but rises in the demand shock case.

Once again, analogous results to expressions (36) and (37) can be derived for the case of aggregate investment. In this case, however, the approximation is based on the idea that a profit-maximizing firm equates the marginal product of investment to the user cost of capital for the various investment goods. Thus, the weights in the approximation are proportional to the marginal return to investment rather than to marginal utility. Nevertheless, the intuition is the same: an increase in demand for a capital good associated with a price decline implies a decline in marginal product; while an increase in quantity resulting from a change in the production process that makes the good more useful results in an increase in marginal product.

7 Conclusion

This paper points out that allowing for demand shocks and non-homothetic preferences has important implications for the well-known concept of "consumer substitution bias". Furthermore, the empirical results presented in Section 3 show that the standard "consumer substitution bias" story does not provide an adequate explanation for differences among alternative empirical quantity indexes. Demand shocks have the potential to explain some of these "anomalies".

In particular, the usual story regarding the relationship between fixed-weighted Laspeyres and chain-weighted Fisher indexes breaks down in the presence of demand shocks. Namely, Fisher quantity index may grow *either* more slowly or more quickly than the fixed-weighted Laspeyres index in the periods following the base period. Moreover, chain bias is likely to be more severe in the presence of demand shocks.

The interpretation of real GDP statistics, however, also becomes more complicated in the presence of demand shocks and non-homothetic preferences (or aggregators that lack the linear homogeneity property). Namely, the results deriving "superlative" index numbers from their theoretical aggregates such as the Konus index break down.

These results suggest that further empirical work is needed to fully to understand the nature of differences among alternative empirical quantity indexes. Moreover, on a theoretical level, the implications of more general preference specifications for standard index number theory results seem like a worthwhile topic for future research.

A A Simple Numerical Example

Consider a simple numerical example of the apple and orange economy. The labor force consists of two yeoman farmers. Farmer 1 produces good 1 (oranges) and farmer 2 produces good 2 (apples). Farmer i's utility is,

$$U_i = C_i - \frac{1}{\gamma} L_i^{\gamma}, \tag{39}$$

where C_i is a consumption aggregate, L_i is the labor supply of farmer i, and γ is a parameter determining the elasticity of labor supply. The consumption aggregate takes the constant elasticity of substitution (CES) form,

$$C_i = [z_1^{\frac{1}{\eta}} c_{1,i}^{\frac{\eta-1}{\eta}} + z_2^{\frac{1}{\eta}} c_{2,i}^{\frac{\eta-1}{\eta}}]^{\frac{\eta}{\eta-1}}, \tag{40}$$

where z_j is an exogenous demand shock to good j, $c_{j,i}$ is consumption of good j by farmer i, and η is a parameter describing the substitutibility of the consumption goods.[29] I assume standard values of the parameters, $\gamma = 2$ and $\eta = 3$. Notice that the demand shocks differ across goods, but are the same for both farmers. To simplify the notation, set $z_1 = 1$. A little algebra shows that the farmers' demand for apples and oranges are given by,

$$c_{1,i} \quad = \frac{I}{1+z_2 p_2^{1-\eta}}, \tag{41}$$

$$c_{2,i} \quad = \frac{I}{1+z_2 p_2^{1-\eta}} \left[\frac{z_2}{p_2^{\eta}}\right], \tag{42}$$

where I is nominal income and p_j is the price of good j. Normalize $p_1 = 1$. The product of the price and consumption indexes is nominal income, $PC = I$. Thus, the farmers' utility can be rewritten as,

$$U_i = \frac{I_i}{P} - \frac{1}{\gamma} L_i^{\gamma}. \tag{43}$$

The yeoman farmer's nominal income equals his sales revenue,

$$I_i = p_i y_i. \tag{44}$$

The production function is linear,

$$Y_j = A_j L_j, \tag{45}$$

where A_j reflects the labor productivity of farmer j. Substituting the expression for nominal income (44) and the production function (45) into the utility function (43), we have,

$$U_i = \frac{p_i A_i L_i}{P} - \frac{1}{\gamma} L_i^{\gamma}. \tag{46}$$

[29] Another paper that analyzes index numbers for the CES functional form is Feenstra and Reinsdorf (2003).

The farmer's utility-maximizing labor supply can be found by maximizing equation (46) with respect to his labor supply, which gives,

$$L_i = (\frac{A_i p_i}{P})^{\frac{1}{\gamma-1}}. \tag{47}$$

Finally, by setting output equal to the combined demands of the two farmers, one can show that the market-clearing price is,

$$p_2 = \left[\frac{z_2}{(\frac{A_2}{A_1})^{\frac{\gamma}{\gamma-1}}}\right]^{\frac{1}{\frac{1}{\gamma-1}+\eta}}. \tag{48}$$

The expression for p_2 shows that if apple production technology improves (A_2 increases), apple prices fall, all else constant. However, if society's taste for apples increases (z_2 increases), then apple prices rise, all else constant.

Suppose there are two periods. In the first period, $A_1 = A_2 = 1$, $z_1 = z_2 = 1$ and $p_1 = p_2 = 1$. By symmetry, equal amounts of the two goods are produced: $y_1 = y_2 = 1.41$. In period 2, a demand or supply shock occurs. The shocks have the following characteristics: the demand shock increases z_2 from 1 to 2, and the supply shock increases A_1 by 3.5% and A_2 by 16%. Following the demand shock, output increases to $y_1 = 1.554$ and $y_2 = 1.848$, while the price of good 2 rises to 1.19. Following the supply shock, output increases to $y_1 = 1.560$ and $y_2 = 1.851$, while the price of good 2 falls to 0.94.

Let us now consider the national accounts statistics for this economy. The Laspeyres index rises by the same amount in either the demand shock or the supply shock case. However, in the demand shock case, $Q_F > Q_L$ since the relative price of good 2 rises, while in the supply shock case, $Q_F < Q_P$ since the relative price of good 2 falls.

B Proofs of Propositions

Proof of Proposition 1a

Consider the minimization problem,

$$C(u_t, p_t) = \min_q \{p_t.q_t : F(q_t) \geq F_t\}, \tag{49}$$

where F_t is a linearly homogenous function, q_t is the vector of period t quantities, and p_t is the vector of period t prices. Let q_t reflect quantities that have been chosen to solve this optimization problem. A standard result shows that $C(u_t, p_t)$ is concave in prices. Moreover, the assumption that F_t is a linearly homogenous implies that $C(u_t, p_t)$ is separable in prices: that is, $C(u_t, p_t) = f(p_t)g(u_t)$.

By the concavity of C,

$$C(u_0, p_1) \leq C(u_0, p_0) + \nabla_p C(u_0, p_0)(p_1 - p_0) \tag{50}$$
$$= C(u_0, p_0) + q_0(p_1 - p_0) \tag{51}$$
$$= p_1 q_0, \tag{52}$$
$$\tag{53}$$

where the second equality follows from the envelope theorem. By definition,

$$C(u_1, p_1) = p_1 q_1. \tag{54}$$

Dividing expression (54) by expression (53) and using the separability assumption we have,

$$\frac{p_1 q_1}{p_1 q_0} \leq \frac{C(u_1, p_1)}{C(u_0, p_1)} \tag{55}$$
$$= \frac{g(u_1)}{g(u_0)}. \tag{56}$$

By an analagous argument, one can show that

$$\frac{p_0 q_1}{p_0 q_1} \geq \frac{g(u_1)}{g(u_0)}. \tag{57}$$

Putting these expressions together and recalling the definitions of Q_L, Q_F, and Q_P, we have the result,

$$Q_P \leq Q_F \leq Q_L. \tag{58}$$

Now it suffices to show that the optimization problem described by (49) follows from from rational optimizing behavior. For the purposes of this proposition, we assume that F_t is a consumption aggregator, such as the Dixit-Stiglitz index commonly used in macroeconomics. A standard result shows that given a linearly homogenous consumption aggregator function, the decision of how much of the consumption aggregate F to consume, and the decision of how to divide the aggregate among goods can be solved separately. Thus, the described by (49) simply follows from the household's optimal consumption problem.

The key point here is that since the consumption aggregator F is stable, changes in price arise from factors outside the household. Actually, for this purpose, it is possible to allow for multiplicative shifts in F i.e. $F_t = A_t F$.

Proof of Proposition 2a

The same approach can essentially be applied to the case of a firm's capital expenditure decision. Suppose the firm's production function, $G(L_t, K_t)$ depends on a capital aggregate K_t. Let us now make the strong assumption that $K_{t+1} = K_t(1-\delta)+I_t$, where I_t is a linearly homogenous investment aggregate and δ is the depreciation rate. Moreover, p_t reflects the rental rate for capital services of a particular type, ρq_t, where ρ is the gross rate of return on capital and q_t is the nominal value of the capital good. Given these assumptions, Proposition 2 follows directly from setting F_t equal to the investment aggregator function I_t.

Once again, the key point here is that the fluctuations in factor prices arise from factors outside the firm "demanding" the factor inputs.

Proof of Proposition 3a

Suppose that the production possibilities set for the economy can be represented as,

$$S = (y, x) : g(y) = f(x), \tag{59}$$

where y is a vector of outputs and x is a vector of inputs. Also assume that $g(y)$ is linearly homogeneous in y, which implies that the scale of production $g(y)$ does not affect the output mix. Clearly, these are strong assumptions. However, they are necessary to provide a simple justification for the output aggregator $g(y)$.

Given a particular level of inputs x, the firm must decide which output to produce. The firm chooses outputs to maximize revenue,

$$R(p, f(x)) = \max_y \{p.y : g(y) = f(x)\}, \tag{60}$$

subject to the constraint that it must produce within the production possibilities frontier. The structure of S implies that the problem defined by equation (60) can be solved separately from the problem of how much aggregate output $g(y)$ to produce.

Moreover, the assumption that g is linearly homogeneous implies that,

$$R(p, f(x)) = r(p)f(x), \tag{61}$$

where $r(p)$ is the unit revenue function $R(p, 1)$. To see this, note that at a given price vector, the firm always has the option of scaling up its production by, say, a factor of k and thus scaling up its revenue by k. Consequently, its revene at various levels of output must be proportional if it is maximizing profits.

By the convexity of the revenue function,

$$R(p_1, f(x_0)) \geq R(p_0, f(x_0)) + \nabla_p R(p_0, f(x_0))(p_1 - p_0) \tag{62}$$
$$= R(p_0, f(x_0)) + y_0(p_1 - p_0) \tag{63}$$
$$= p_1 y_0, \tag{64}$$
$$\tag{65}$$

where the second equality follows from the envelope theorem. By definition,

$$R(p_1, f(x_1) = p_1 q_1. \tag{66}$$

Dividing expression (66) by expression (65) and using the separability assumption we have,

$$\frac{p_1 q_1}{p_1 q_0} \geq \frac{R(p_1, f(x_1))}{R(p_1, f(x_0)} \tag{67}$$
$$= \frac{g(y_1)}{g(y_0)}. \tag{68}$$

An analagous argument shows,

$$\frac{p_0 q_1}{p_0 q_1} \leq \frac{g(u_1)}{g(u_0)}. \tag{69}$$

Therefore,

$$Q_{L,1} \leq Q_{F,1} \leq Q_{P,1}, \tag{70}$$

where the inequalities are strict if the profit function has a unique optimum.

Proof of Proposition 4a

Proposition 4 follows directly from Proposition 3, taking the final demanders to be firms.

Proofs of Propositions 1b,2b (converse results)

So far, I have shown that supply shocks imply that $Q_P < Q_F < Q_L$ while demand shocks imply that $Q_L < Q_F < Q_P$. What remains to show of the proposition is a sort of converse result. Namely, $Q_P < Q_F < Q_L$ is consistent with homothetic utility maximization, while $Q_L < Q_F < Q_P$ is consistent with profit maximization with a homogenous output aggregator g.

I make use of the Diewert-Afriat (1973) recoverability results to prove these results. Let us first normalize the prices in each period such that,

$$p_k q_k = 1. \tag{71}$$

Define the matrix of Afriat cross-coefficients as,

$$D_{j,k} = p_j q_k - 1 \qquad \text{for} \qquad 1 \leq j,k \leq K. \tag{72}$$

Now consider the following system of inequalities involving K unknown variables $\{\lambda_k\}$,

$$\lambda_j D_{j,k} \geq \lambda_k - \lambda_j \qquad \text{for} \qquad 1 \leq k \leq K, j \neq k \tag{73}$$

$$\lambda_k \geq 1 \qquad \text{for} \qquad k = 1, ..., K, \tag{74}$$

If a solution to the system of inequalitities (73) and (**??**) exists, say λ_k^* for $k = 1, , K$ then a utility function $h(q)$ which rationalizes the K observation data set can be defined as follows:

$$h(q) \equiv \min_k \{\lambda_k * p^k q : k = 1, ..., K\}. \tag{75}$$

If $K = 2$ the system of inequalities (73) and (74) can be rewritten as,

$$\lambda_1 D_{1,2} \geq \lambda_2 - \lambda_1 \tag{76}$$

$$\lambda_2 D_{2,1} \geq \lambda_1 - \lambda_2. \tag{77}$$

The solution to this set of equations can in turn be found by setting $\lambda_1 = 1$ and looking for a $\lambda_2 > 0$ that solves the inequalities,

$$D_{1,2} \geq \lambda_2 - 1 \tag{78}$$

$$\lambda_2 D_{2,1} \geq 1 - \lambda_2. \tag{79}$$

Using definitions (72) and the normalization conditions (71) for $k = 1$, one can then show that the inequality (78) is equivalent to,

$$\frac{p_1 q_2}{p_1 q_1} - 1 \geq \lambda_2 - 1 \tag{80}$$

or

$$\lambda_2 \leq Q_L. \tag{81}$$

Similarly, using the definition (72) and the normalization conditions (71) for k = 2, one can show that the inequality (79) is equivalent to,

$$\lambda_2 [\frac{p_2 q_1}{p_2 q_2}] - 1 \geq 1 - \lambda_2 \tag{82}$$

or

$$[\frac{p_2 q_2}{p_2 q_1}]^{-1} \geq [\lambda_2]^{-1} \tag{83}$$

or

$$\lambda_2 \geq \frac{p_2 q_2}{p_2 q_1} \equiv Q_P \tag{84}$$

Thus, the necessary and sufficient conditions for the existence of a positive λ_2 which satisfies (78) and (79) are,

$$Q_P \leq \lambda_2 \leq Q_L. \tag{85}$$

Proofs of Propositions 3b,4b (converse results)

This proof follows a very similar logic to the proof above for the supply shock case. Consider the following system of inequalities involving K unknown variables $\{\lambda_k\}$,

$$\lambda_j D_{j,k} \leq \lambda_k - \lambda_j \qquad \text{for} \qquad 1 \leq k \leq K, j \neq k \tag{86}$$

$$\lambda_k \geq 1 \qquad \text{for} \qquad k = 1, ..., K, \tag{87}$$

If a solution to the inequalitities (87) exists, say λ_k^* for $k = 1, , K$ then an output aggregator function $g(q)$ which rationalizes the K observation data set can be defined as,

$$g(q) \equiv \max_k \{\lambda_k * p^k q : k = 1, ..., K\}. \tag{88}$$

If K=2, the inequalities (87) can be rewritten as,

$$\lambda_1 D_{1,2} \leq \lambda_2 - \lambda_1 \tag{89}$$

$$\lambda_2 D_{2,1} \leq \lambda_1 - \lambda_2. \tag{90}$$

This system of inequalities can in turn be rewritten by setting $\lambda_1^* = 1$ yielding the system,

$$D_{1,2} \leq \lambda_2 - 1 \tag{91}$$

$$\lambda_2 D_{2,1} \leq 1 - \lambda_2. \tag{92}$$

Using the definitions (72) and the normalization condition (71), the inequality (91) can be rewritten as,

$$\frac{p_1 q_2}{p_1 q_1} - 1 \leq \lambda_2 - 1 \tag{93}$$

or

$$\lambda_2 \leq Q_L. \tag{94}$$

Similarly, the inequality (92) can be rewritten as,

$$\lambda_2 [\frac{p_2 q_1}{p_2 q_2}] - 1 \leq 1 - \lambda_2 \tag{95}$$

or

$$\lambda_2 \leq \frac{p_2 q_2}{p_2 q_1} \equiv Q_P \tag{96}$$

Thus necessary and sufficient conditions for the existence of a positive λ_2 which satisfies (91) and (92) are,

$$Q_L \leq \lambda_2 \leq Q_P. \tag{97}$$

References

, 1988.

Afriat, S.N., "The Construction of Utility Functions from Expenditure Data," *International Economic Review*, 1967, *8*, 67–77.

Aizcorbe, Ana M. and Patrick C. Jackman, "The Commodity Substitution Effect in CPI Data," *Monthly Labor Review*, December 1993, pp. 25–33.

Allen, R.G.D., *Index Numbers in Theory and Practice*, London: Macmillan, 1975.

Alterman, W. F., W. E. Diewert, and R. C. Feenstra, *International Trade Price Indexes and Seasonal Commodities*, Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics, 1999.

Balk, Bert, "Direct and Chained Indicies: A Review of Two Paradigms," June 2004. Working paper http://www.ipeer.ca/papers/Balk,June3,2004Vancouver.pdf.

Balk, Bert M., "Changing Consumer Preferences and the Cost-of-Living Index: Theory and Nonparametric Expressions," *Journal of Economics*, 1989, *50*, 157–169.

Basu, Susanto and John G. Fernald, "Returns to Scale in U.S. Production: Estimates and Implications," *The Journal of Political Economy*, April 1997, *105* (2), 249–283.

__ **and** __ , "Aggregate Productivity and Aggregate Technology," May 2001. Working Paper.

BEA, United States, "The Comprehensive Revision of the U.S. National Income and Product Accounts: A Review of Revisions and Major Statistical Changes," *Survey of Current Business*, 1991. http://www.bea.gov/bea/an1.htm.

__ , "U.S. National Income and Product Accounts: Annual and Quarterly Chain Measures of Quantity and Prices," 1998. OECD Meeting of National Accounts Experts, http://www1.oecd.org/std/nameet98/27BEA2.PDF.

Bortkiewicz, L.v., "Zweck und Stuktur einser Preinsindexzahl, Nordosk Statistik Tidskrift," in R.G.D. Allen, ed., *Index Numbers in Theory and Practice*, 1975, pp. 62–65.

Braithwait, Steven D., "The Substitution Bias of the Laspeyres Price Index: An Analysix Using Estimated Cost-of-Living Indexes," *American Economic Review*, March 1980, (70), 64–77.

Croushore, Dean and Tom Stark, "A Real-Time Data Set for Macroeconomists," *Journal of Econometrics*, November 2001, *105*, 111–130.

Diewert, W. E., "Afriat and Revealed Preference Theory," *The Review of Economic Studies*, 1973, *40*, 419–425.

__ , "Functional Forms for Profit and Transformation Functions," *Journal of Economic Theory*, 1973, *6* (3), 284–316.

__ , "Functional Forms for Revenue and Factor Requirement Functions," *International Economic Review*, 1974, *15*, 119–130.

__ , "Exact and Superlative Index Numbers," *Journal of Econometrics*, May 1976, (4), 115–146. and reprinted as Chapter 8 in Diewert and Nakamura (1993, pp. 223-252).

__ , "Superlative Index Numbers and Consistency in Aggregation," *Econometrica*, 1978, *46*, 883–900. and reprinted as Chapter 9 in Diewert and Nakamura (1993, pp. 223-252).

__ , "The Test Approach to Economic Comparisons," in "Measurement in Economics" 1988, pp. 67–86.

__ , "Fisher Ideal Output, Input and Productivity Indexes Revisited," *The Journal of Productivity Analysis*, 1992, (3), 211–248. and reprinted as Chapter 13 in Diewert and Nakamura (1993, pp.211-248).

__ , "Index Numbers," in M. Milgate In J. Eatwell and P. Newman, eds., *The New Palgrave: A Dictionary of Economics Vol. 2*, London: Macmillan, 1992, pp. 767–780. and reprinted as Chapter 5 in Diewert and Nakamura (1993, pp. 71-104).

__ , "Index number Issues in the Consumer Price Index," *The Journal of Economic Perspectives*, 1998, *12* (1), 47–58.

__ , "Harmonized Indexes of Consumer Prices: Their Conceptual Foundations," January 2002a. Working Paper http://www.econ.ubc.ca/diewert/other.htm.

__ **and A. O. Nakamura**, "The Measurement of Aggregate Total Factor Productivity Growth," *Handbook of Econometrics*, April 2002. Forthcoming in Elsevier Science *Handbook of Econometrics*, Volume 6, edited by J.J. Heckman and E.E. Leamer.

__ **and K. Fox**, "Can Measurement Error Explain the Productivity Paradox?," *Canadian Journal of Economics*, 1999, *32* (2), 251–281.

__ **and T. J. Wales**, "Flexible Functional Forms and Tests of Homogeneous Separability," *Journal of Econometrics*, 1995, *67*, 259–302.

Dixit, A.K. and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 1977, *67*, 297–308.

Dynan, Karen E. and Douglas W. Elmendorf, "Do Provisional Estimates of Output Miss Economic Turning Points," September 2001. Working Paper.

Eichhorn, W., “Fisher's Tests Revisited,” *Econometrica*, 1976, *44*, 247–256.

__ **and J. Voeller**, *Theory of the Price Index*, Berlin: Springer-Verlag, 1976. Lecture Notes in Economics and Mathematical Systems 140.

Faust, Jon, John H. Rogers, and Jonathan Wright, “News or Noise in G-7 GDP Announcements,” August 2001. Working Paper.

Feenstra, Robert C., “Exact Hedonic Price Indexes,” *Review of Economics and Statistics*, 1995, *78* (4), 634–653.

__ **and Marshall B. Reinsdorf**, “Should Exact Index Numbers have Standard Errors? Theory and Application to Asian Growth,” December 2003. NBER Working Paper 10197.

__ **and Marshall Reinsdorf**, “An Exact Price Index for the Almost Ideal Demand System,” *Economics Letters*, 2000, *66* (2), 159–162.

__ **and Matthew D. Shapiro**, “High-Frequency Substitution and the Measurement of Price Indexes,” March 2001. NBER Working Paper 8176.

Fisher, Franklin M., “Aggregate Production Functions and the Explanation of Wages: A Simulation Experiment,” *The Review of Economics and Statistics*, November 1971, *LIII* (4), 305–325.

Fisher, Irving, *The Making of Index Numbers*, Boston: Houghton Mifflin, 1922.

Forsyth, F. and R. Fowler, “The Theory and Practice of Chain Price Index Numbers,” *The Journal of the Royal Statistical Society, Series A*, 1981, *144* (2).

Gollop, Frank M. and Dale W. Jorgenson, “Sectoral Measures of Labor Cost for the United States, 1948-1978,” in Jack E. Triplett, ed., *The Measurement of Labor Cost*, Vol. 44 of NBER Studies in Income and Wealth, Chicago, IL: University of Chicago Press, 1983, pp. 185–235 and 503–520.

Gordon, Robert J., “Has the ‘New Economy’ Rendered the Productivity Slowdown Obsolete,” June 1999. Working Paper.

Jackson, Chris, “The Effect of Rebasing on GDP,” Technical Report, Statistics Canada Second Quarter 1996. Catalogue no. 13-001-XPB.

Jorgenson, Dale W. and Barbara M. Fraumeni, “The Output of the Education Sector,” in Zvi Griliches, ed., *Output Measurement in the Services Sector*, Chicago, IL: University of Chicago Press, 1992.

__ **, ed.**, *Productivity*, Vol. 1, Cambridge: MIT Press, 1995.

__ **, Frank Gollop, and Barbara Fraumeni**, *Productivity and U.S. Economic*

Growth, Cambridge: Harvard University Press, 1987.

Koenig, Evan F., Sheila Dolmas, and Jeremy Piger, "The Use and Abuse of "Real-time" Data in Economic Forecasting," August 2001. Working Paper.

Konyus, A.A. and S.S. Byushgens, *K probleme pokupatelnoi cili deneg, Voprosi Konyunkturi II: 1, 151-172* 1926. English Title: On the problem of the purchasing power of money.

Landefeld, Steven J. and Robert P. Parker, "Preview of the Comprehensive Revision of the National Income and Product Accouts: BEA's New Featured Measures of Output and Prices," *Survey of Current Business*, July 1995, pp. 31–38. http://www.bea.gov/bea/an1.htm.

__ **and** __ , "BEA's Chain Indexes, Times Series, and Measures of Economic Growth," *Survey of Current Business*, May 1997, pp. 58–68. http://www.bea.gov/bea/an1.htm.

Lucas, Robert E. Jr., "Expectations and the Neutrality of Money," *Journal of Economic Theory*, April 1972, *4*, 103–124.

Mankiw, N. Gregory and Matthew D. Shapiro, "News or Noise: An Analysis of GNP Revisions," *Survey of Current Business*, May 1986, pp. 20–25.

Manser, Marilyn E. and Richard J. McDonald, "An Analysis of Substitution Bias in Measuring Inflation, 1959-85," *Econometrica*, July 1988, *56* (4), 909–930.

Orphanides, Athanasios, "Monetary Rules Based on Real-Time Data," *American Economic Review*, 2001.

Parker, Robert P. and Jack E. Triplett, "Chain-Type Measures of Real Output and Prices in the U.S. National Income and Product Accounts: An Update," Technical Report, Bureau of Economic Analysis 1996. http://www.unece.org/stats/documents/ocde/std/na/rd.97/10.e.doc.

Reinsdorf, Marshall, "Formula Bias and Within-Stratum Substitution Bias in the US CPI," *Review of Economics and Statistics*, 1998, *8* (2), 175–187.

Romer, David, *Advanced Macroeconomics*, New York: McGraw Hill, 2001.

Roubini, Nouriel, "Productivity Growth, Its Slowdown in the 1973-90 Period and its Resurgence in the 1990s: Truth or a Statistical Fluke," 1998. Internet Article http://pages.stern.nyu.edu/ nroubini/MEASURE.HTM.

Shapiro, Matthew D. and David W. Wilcox, "Alternative Strategies for Aggregating Prices in the CPI," *Federal Reserve Bank of St. Louis Review*, May/June 1997, pp. 113–125. Also available as NBER Working Paper 5980.

Solow, Robert M., "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics*, August 1957, *39* (3), 312–320.

Syverson, Chad, "Product Substitutability and Productivity Dispersion," July 2003. Working Paper.

Szulc, Bohdan, "Linking Price Index Numbers," in W. Erwin Diewert and Claude Montmarquette, eds., *Price Level Measurement: Proceedings from a conference sponsored by Statistics Canada*, Ottawa: Minister of Supply and Services Canada, 1983, pp. 537–566.

Tevlin, Stacey and Karl Whelan, "Explaining the Investment Boom of the 1990s," March 2000. Working Paper.

Tornqvist, Leo, "The Bank of Finland Consumption Price Index," *Bank of Finland Monthly Bulletin*, 1936, (10), 1–8.

Triplett, Jack E., "Economic Theory and BEA's Alternative Quantity and Price Indexes," *Survey of Current Business*, April 1992, pp. 49–52. http://www.bea.gov/bea/an1.htm.

Young, Allan H., "Alternative Measures of Change in Real Output and Prices," *Survey of Current Business*, April 1992, pp. 32–48. http://www.bea.gov/bea/an1.htm.

__ , "Alternative Measures of Change in Real Output and Prices, Quarterly Estimates for 1959-92," *Survey of Current Business*, March 1993, pp. 31–41. http://www.bea.gov/bea/an1.htm.

Table 1:

Average Labor Productivity Growth Rate under different Weighting Regimes		
	Fixed-1987 Weighted Measure	**Chain-weighted Measure**
1950's	2.07	2.07
60's	2.37	2.80
70's	1.80	2.24
80's	0.69	1.02
90-95	1.73	1.24

Table 2

A Comparison of Fixed-Weighted Laspeyres and Chain-Weighted Fisher Indexes for 5 Years following the Base Year

	Aggregate Investment		Consumption		Government Expenditure	
Base Year	Laspeyres	Fisher	Laspeyres	Fisher	Laspeyres	Fisher
1958	1.36	1.88	3.58	3.45	3.95	3.55
1972	3.27	4.23	3.57	3.60	1.04	0.70
1982	1.59	0.93	4.00	4.19	3.86	3.73
1987	0.43	0.03	1.82	1.98	1.56	1.77

Table 3

The Difference between the Fixed-Weighted Laspeyres and Chain-Weighted Fisher Indexes for the Year Before and After the Base Year

	Aggregate Investment		Consumption		Government Expenditure		Imports	
Base Year	Before	After	Before	After	Before	After	Before	After
1958	-0.54	-1.58	-0.13	0.31	0.53	-4.66	-0.53	3.69
1972	-1.10	-0.42	-0.21	-0.73	1.11	0.05	0.14	1.91
1982	-4.47	0.58	0.39	-0.55	-0.21	0.19	0.38	-4.92
1987	0.04	2.80	-0.20	-0.31	0.40	-0.47	-1.34	-0.15

Table 4

A Comparison of Laspeyres and Paasche Indexes

	Aggregate Investment		Consumption		Government Expenditure		Imports	
Period	Laspeyres	Paasche	Laspeyres	Paasche	Laspeyres	Paasche	Laspeyres	Paasche
1958q1-1972q4	4.30	4.50	4.27	4.07	3.14	2.99	7.67	7.72
1972q1-1982q4	3.25	2.81	2.58	2.23	1.78	1.55	3.53	2.38
1982q1-1987q4	2.57	1.66	4.04	3.90	3.69	3.95	13.03	11.76

Table 5

A Comparison of Chained Fisher and Direct Fisher Indexes

	Aggregate Investment		Consumption		Government Expenditure	
Period	direct	chained	direct	chained	direct	chained
1958q1-1972q4	4.40	5.12	4.17	4.14	3.27	2.79
1972q1-1982q4	3.02	4.06	2.39	2.24	1.42	1.27
1982q1-1987q4	2.07	1.61	3.97	4.73	4.12	4.15

Figure 1

A Comparison of 5-Year Moving Average Growth Rates for Real GDP:
Fixed-weighted indexes minus Chain-weighted index

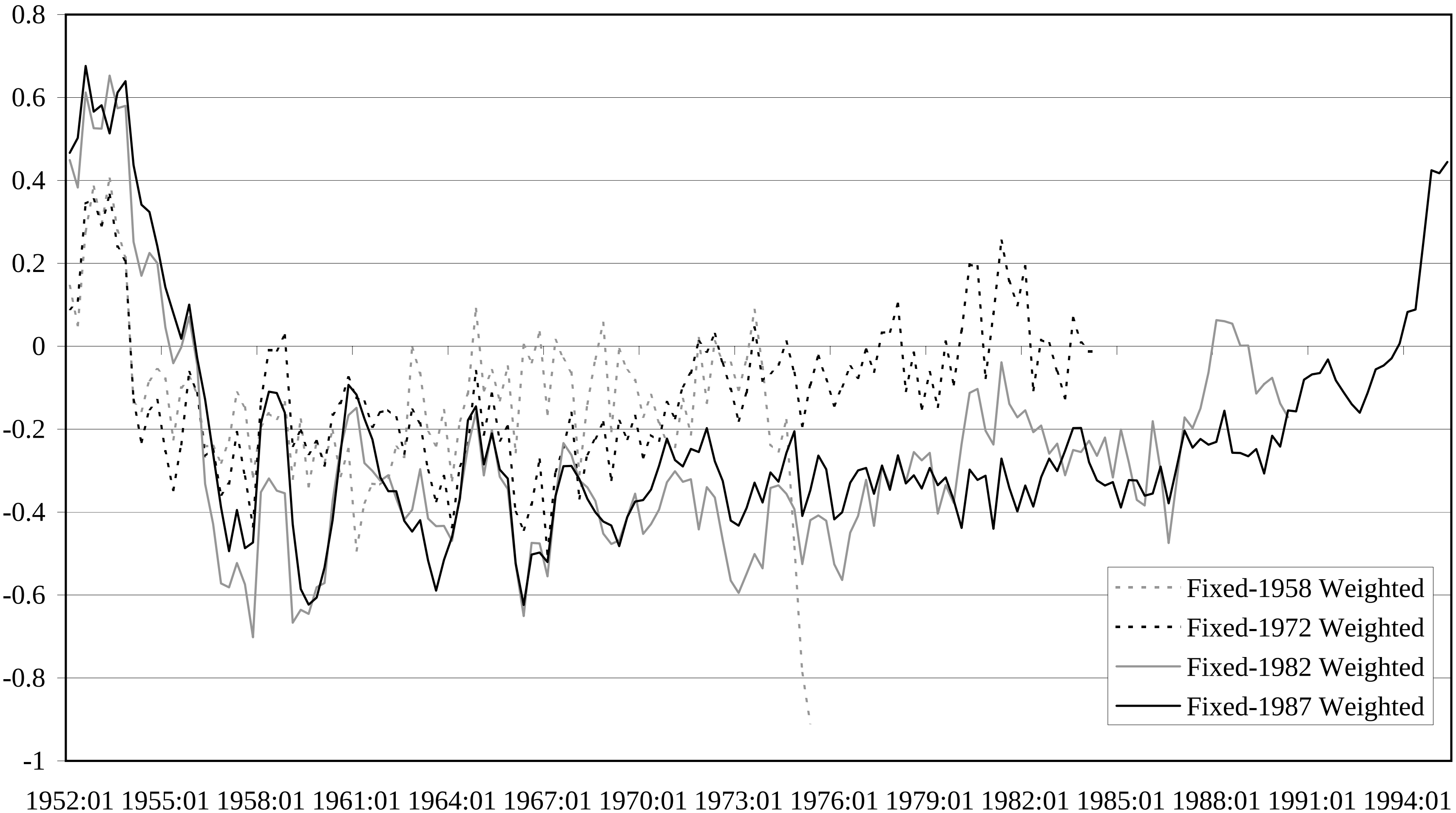

Figure 2

A Comparison of 5-Year Moving Average Growth Rates for Aggregate Investment: Fixed-weighted series minus Chain-weighted series

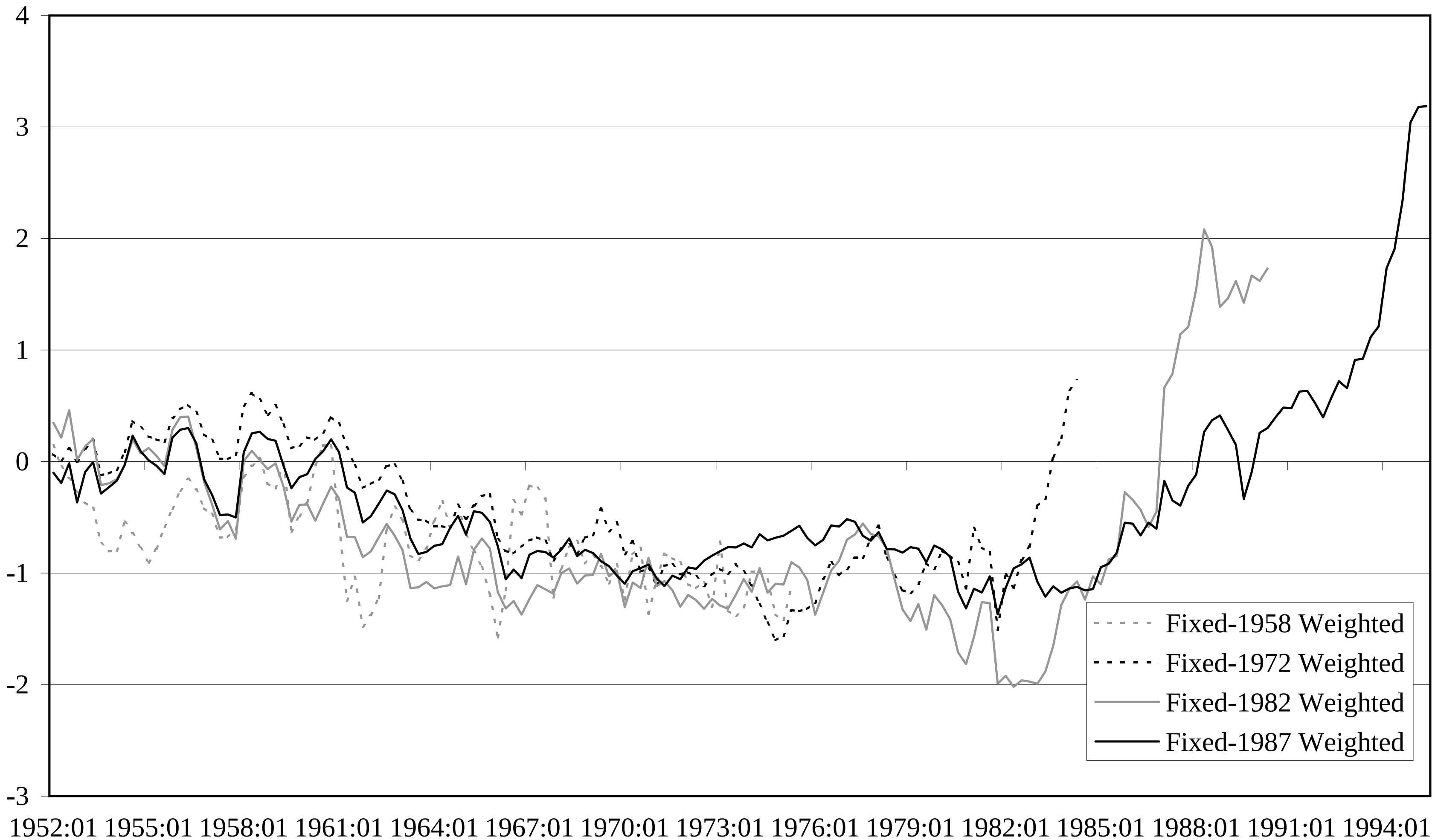

Figure 3

A Comparison of 5-Year Moving Average Growth Rates for Government Expenditure: Fixed-weighted series minus Chain-weighted series

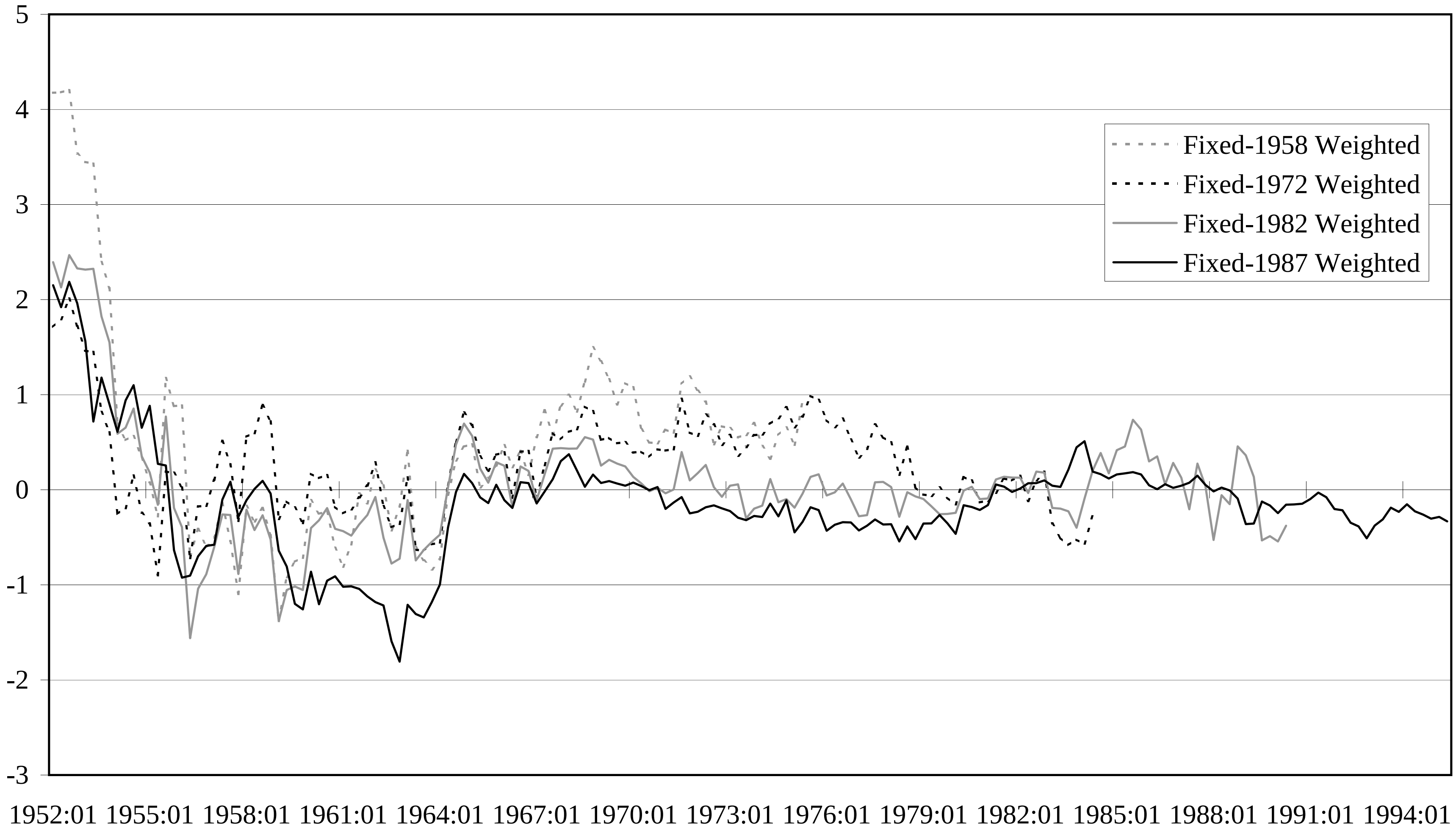

Figure 4

A Comparison of 5-year Moving Average Growth Rates for Aggregate Consumption: Fixed-weighted measures minus chain-weighted measure

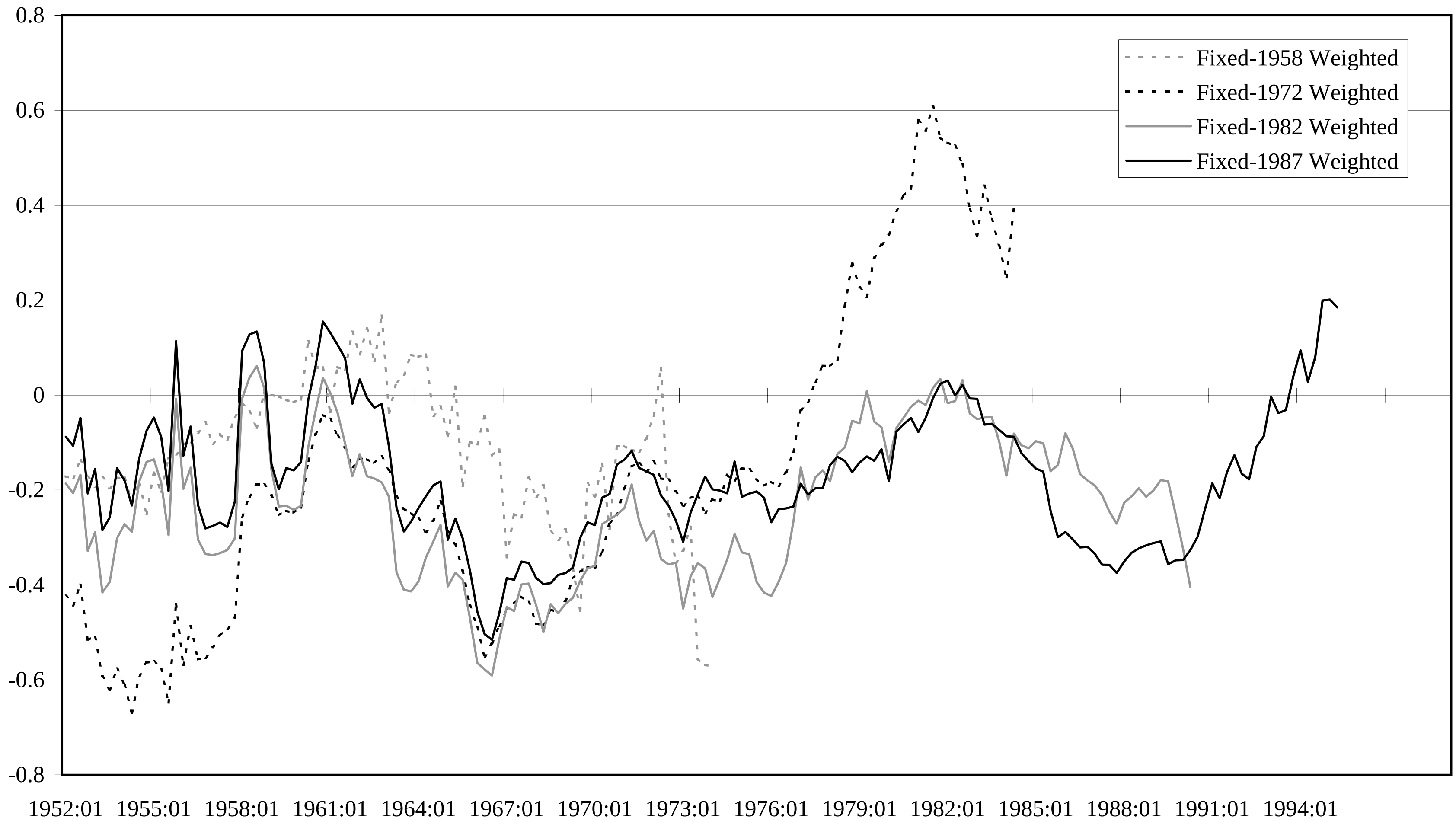

Figure 5:
Effects of a Supply Shock on Konus Quantity Index and Empirical Quantity Indexes

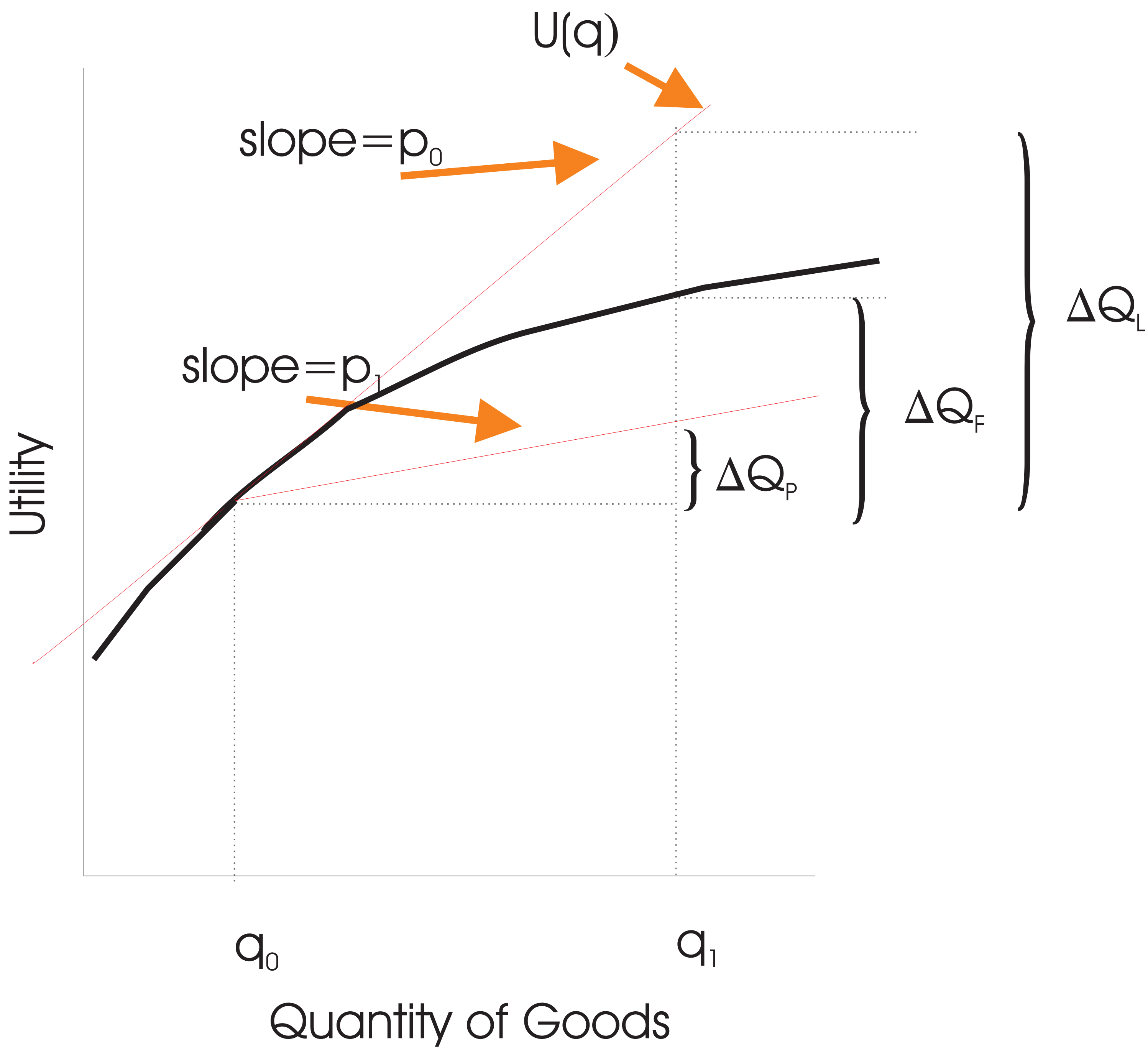

Figure 6:
Effects of a Demand Shock on Konus Quantity Index and Empirical Quantity Indexes