A Critique of Antitrust Econometrics: Aggregation, the Representative Consumer, and the Broader Concerns of the New Brandeis School*

Gabriel A. Lozada

November 11, 2021

Abstract. Some economists argue antitrust policy should be based on empirical methods used by the Industrial Organization subdiscipline of economics, but those methods contain assumptions that non-economists should recognize. Those assumptions underlie econometric "identification," and underlie treating aggregate demand as if it were generated by a representative consumer (Muellbauer's "generalized linear" preferences). We explain aggregation bias in Almost Ideal Demand System models, then show that data limitations make it even harder to justify economists' restricting aggregate demands as one would the demand of one individual. Such problems notwithstanding, the main problem with antitrust econometrics may be that there is not enough of it. Whether firms maximize profit is under-studied empirically; many may maximize return on assets instead, leading to firms with assets and employees below their profit-maximizing level. There is insufficient empirical study of this and many other topics of concern to New Brandeisians.

Keywords: Antitrust econometrics; Almost Ideal Demand System (AIDS); Myth of Shareholder Value; Financialization of Management; New Brandeis School

*Department of Economics, University of Utah, Salt Lake City, UT 84112, USA; lozada @economics.utah.edu; 801-581-7650. I would like to thank Mark Glick for suggesting this topic, and thank him and editor Bill Curran for making helpful criticisms on earlier drafts. This work has been partially supported by the Institute for New Economic Thinking.

In a recent Antitrust magazine article, Carl Shapiro[1] makes the case for what he calls the "modern" approach to the future of antitrust. He stakes out his position in opposition to the "consumer welfare" camp and the "populist" camp—by the latter he means the New Brandeis School. Under Shapiro's "modern" approach, antitrust is reduced to an "effects" analysis based on modern Industrial Organization. At the apex of this approach is empirical Industrial Organization, where econometrics is employed to quantify anticompetitive effects (negative welfare effects) and efficiencies (positive welfare effects). This exercise usually involves I.O. economists on both sides of the case. Because of the technical nature of the practice, the economists take center stage. However, many economists on both sides have been trained to accept certain assumptions, as well as to limit the scope of their inquiries in important ways.

In Baker and Rubinfeld's[2] survey article, while problems with "identification[3]" and "functional forms[4]" are discussed (and will be revisited and explained by us below in Section 1), no mention is made of the strong assumptions needed in order to justify economists' common practice of applying restrictions from economic theory to the estimation of aggregate demand systems. Sections 2 through 5 of this paper is devoted to explaining what those assumptions are, including new results along those lines. An additional assumption underlying most (though not all) econometric approaches is that income but not wealth determines consumption. That makes little sense outside of a purely theoretical static framework, but abandoning that assumption would require economists to have data not only on household income but also on household wealth, and the latter is often not available. Once all the assumptions needed to justify the econometrics are understood, they need not be accepted, raising questions about how accurate the econometric analyses really are.

Section 6 discusses problems using consumer surplus as a measure of consumer welfare, problems which turn out to be of a similar form to those involved in working with demand systems.

Moving to less mathematical topics, Section 7 questions whether firms maximize profit, which is the assumption underlying almost all neoclassical analysis of firm behavior, and calls for more econometric analysis of this

[1] Carl Shapiro, *Antitrust: What Went Wrong and How to Fix It*, 35/3 ANTITRUST 33 (2021) at 33–34.

[2] Jonathan B. Baker and Daniel L. Rubinfeld, *Empirical Methods in Antitrust Litigation: Review and Critique*, 1/1 AMERICAN LAW AND ECONOMICS REVIEW 386 (1999).

[3] *Id.* at 408.

[4] *Id.* at 413.

question. This Section reviews the arguments that in the immediate post-World War II period there were important firms which willingly sacrificed short-term profit in order to benefit society as a whole, and that some such firms still exist. If true, that behavior should be taken into account before allowing such firms to be subject to a hostile takeover. This Section also points out that Wall Street, and, accordingly, many large U.S. firms, take as firm goals not profit but rather various financial ratios, especially return on assets. Such a perverted goal can be an immediate cause of firms wishing to have as few employees and assets as possible. A merger instigated by managers with such goals can result in merged companies soon having *fewer*, not more, assets and employees than the former companies which merged. This decrease in "size" (measured in assets and employees, rather than in market share or revenues) will falsely be characterized by the pro-merger parties as "efficiencies" when, in fact, it is probably a welfare-reducing, even profit-reducing, aspect of the merger. More empirical analysis of firm motives is needed to help enforcement agencies more carefully distinguish between true merger-specific efficiencies and predatory behavior motivated by short-run financial interests.

The last section of the paper, Section 8, presses further along the lines of critiquing not what econometricians have done but what they have not done. These errors of omission may in the final analysis be a more serious problem than the Industrial Organization econometricians' errors of commission. The problem with elevating modern I.O. to the center of antitrust policy is that if important policy goals do not appear in the I.O. literature they are dismissed as not relevant to true competition concerns. For example, the effects of practices such as mergers on workers, small business, the income distribution, and the environment are ignored. Econometrics in antitrust should be expanded to include analysis from labor economics, environmental economics, urban and regional economics, and other economics subdisciplines, without which the econometrics paints only a misleadingly partial picture.

1. The Subjectivity of Econometrics

Computation of demand and supply curves is probably the most common empirical analysis in antitrust litigation—for example, it is needed in order to estimate how the consumption of one commodity will change when the price of another commodity increases. It turns out that computation of demand and supply curves cannot be done in an a-theoretical, purely empirical way, because those curves are not obvious from market data. Typical

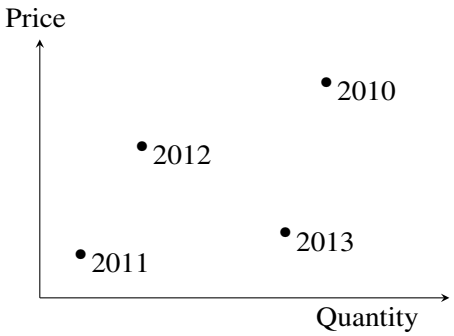

Figure 1. A plot of price and quantity sold for a hypothetical commodity in the years 2010–2013.

observations of market data for price and quantity look like Figure 1, which presents itself as a meaningless scatter of points. This is why the idea of demand curves was unknown to such important and insightful early economists as Adam Smith and David Ricardo, with the first use of demand curves being by Antoine-Augustin Cournot in 1838.[5]

Economists make sense of this situation by imposing *non*-empirically-grounded ideas from economic *theory*. However, these ideas may be inappropriate and lead to faulty policy analysis by giving a misleading impression of how reliable the economic analysis actually is. Non-economists in antitrust should inquire about these assumptions and how justifiable they are.

As an example of how these non-empirical, theoretical assumptions work, an economist looking at Figure 1 may require that the data points are intersections between the demand and supply curves of competitive agents; that demand curves are downward-sloping; and that supply curves are upward-sloping. This enables the economist to rule out one demand curve passing through the data points 2011 and 2012, or 2011 and 2010, or 2011 and 2013, or 2012 and 2010, or 2013 and 2010. If some important non-price determinant of demand, such as income, was similar in 2012 and in 2013, but different in 2011 and in 2010, support increases for one demand curve passing through the 2012 and 2013 points, but being different for 2011 and for 2010. Assuming that supply curves are upward-sloping

[5]Thomas M. Humphrey, *Marshallian Cross Diagrams and Their Uses before Alfred Marshall: The Origins of Supply and Demand Geometry*, 78/March ECONOMIC REVIEW, FEDERAL RESERVE BANK OF RICHMOND 3 (1992) at 3.

narrows down possible locations of supply curves, but still leaves several possibilities: 2011, 2012, and 2010 could all share the same supply curve, or 2011, 2013, and 2010 could all share the same supply curve, or only 2011 and 2010 could share a supply curve and 2012 and 2013 each have their own supply curves; or, each year could have its own supply curve. If this is the market for food and if good or bad weather is known to affect supply, then knowing which years had similar weather would help decide which of these possibilities for supply curves are more likely.

This explains the need for econometricians to impose non-empirically-grounded, theoretical ideas in order to solve what is called the "identification" problem, which refers to this impossibility of using data *alone* to locate the curves which neoclassical economists believe generated the data. Ulrick[6] considers identification to remain a very serious problem an antitrust econometrics. He sharply criticizes the way Hausman, Leonard, and Zona[7] solved their identification problem, and Ulrick writes, "Perhaps the preceding discussion makes demand estimation seem a hopeless cause. Indeed, it is highly doubtful that the conditions that make possible consistent estimation of demand are met." The point is that the question "Is it a demand curve, or is it a supply curve?" (the title of a paper Ulrick cites in his footnote 31) cannot be answered just by collecting data; measurement in economics is more difficult than that, and making even such basic measurements can be controversial.

Econometricians also need to tell their statistical algorithms what sort of general shape to assume for demand curves. For example, are they linear? Parabolic? Or perhaps they are described by a second-degree polynomial, which is "flexible" enough to include the linear and parabolic shapes as special cases. In the same way, the statistical algorithms need to be told what sort of general shape to assume for supply curves. Typically, supply curves are assumed to be upward sloping. This is fine for competitive firms, whose supply curves are their marginal cost curves (above the bottom of average variable cost). Monopolists, however, have no supply curves (they have a supply point), and their marginal cost curves are likely downward-sloping—after all, economies of scale are a common explanation for why a firm was able to become a monopolist in the first place. Therefore, it is not obvious what the best *a priori* assumption to impose on the supply side of a possibly non-competitive market. The upshot is that econometric analysis

[6]Shawn Ulrick, *A Primer on AIDS-Based Models in Antitrust Analysis*, 10/1 EUROPEAN COMPETITION JOURNAL 123 (2014) at 144–5.

[7]Jerry Hausman, Gregory Leonard, and J. Douglas Zona, *Competitive Analysis with Differenciated Products*, 34 ANNALES D'ÉCONOMIE ET DE STATISTIQUE 159 (1994).

of markets requires a non-empirically-grounded theoretical structure, and cannot be better than that structure. In the next several Sections we look at the most commonly-assumed theoretical structure for consumer demand.

2. Econometric Difficulties in Estimating Consumer Demand

From here through Section 6 will be concerned with the market's demand side, where econometricians face another problem besides identification. Muellbauer[8] writes:

> Suppose that n groups of commodities have been defined. The minimum number of parameters necessary to define a demand system, even if the equations have been defined in terms of the $n-1$ independent value shares and after the homogeneity restriction has been imposed, is $(n+2)(n-1)$.

The problem is that in most settings, n is so large that $(n+2)(n-1)$ is a large number. (Imagine how many parameters would be needed to describe how a change in the price of each of 135 different types of breakfast cereals[9] affected the demand for each one of the other cereals.)

If the demand-side data comes from surveys of all the expenditures of a household—as is available from, for example, the U.S. government's Consumer Expenditure Survey[10]—there are very many theoretical relationships that have to hold and that can be used to give structure to the data. This reduces the number of unknowns that have to be estimated, ameliorating or even completely solving the problem that $(n+2)(n-1)$ is a large number.[11] We will call these valuable, theoretically-derived properties of demand curves "(T)" (for "theory"), and postpone for a while detailed description of them.

On the other hand, if the demand-side data is aggregate rather than individual, and/or partial—dealing with only one or a few commodities rather than all commodities—then there is no reason to suppose demand curves obey (T). In fact, in this case surprisingly little can be said about what properties the demand curve should have. This is known as the Sonnenschein-Mantel-Debreu theorem: the market demand curve for a market populated

[8] John Muellbauer, *Aggregation, Income Distribution and Consumer Demand*, 42/4 THE REVIEW OF ECONOMIC STUDIES 525 (1975) at 525.

[9] See Hausman, Leonard, and Zona, *supra* note 7 at 160.

[10] See Wikipedia, *Consumer Expenditure Survey*, available at `https://en.wikipedia.org/wiki/Consumer_Expenditure_Survey`.

[11] These relationships do in turn depend on assuming the households' preferences are unchanging and rational, and could be criticized on those grounds, but such criticism is rarely levelled.

with utility-maximizing rational agents can take the shape of any function that is continuous, has homogeneity of degree zero, and is in accordance with budget balance ("Walras's law").[12] (The second criterion means if all prices double and if all incomes double then the equilibrium quantities are unchanged because what has happened is akin to re-denominating prices from dollars to half-dollars; the third criterion is that consumers must obey their budget constraints.) In particular, market demand curves cannot even be guaranteed to be downward-sloping, and "the problem that $(n+2)(n-1)$ is a large number" looms.

The upshot is that there is very rich theoretical support available for anyone estimating an individual's demand curve, but almost no theoretical support at all for anyone estimating a market demand curve. This is not a problem if one has comprehensive household-level expenditure data: estimate each household's demand function, then add them all up to get the market demand function. But if one does not have comprehensive household-level expenditure data—which is the typical situation in antitrust—then there are many obstacles in the way of finding the market demand curve, and it is important for antitrust practitioners to handle these obstacles correctly.

Economists have responded to the difficulty of estimating the market demand curve by investigating the circumstances under which the market demand curve can be thought of as being derived from "the average consumer," called "exact linear aggregation," or from "a representative (but not mathematically average) consumer," called "exact nonlinear aggregation." We will use "exact aggregation" as an overall term to describe these two types of aggregation. If the circumstances allowing exact aggregation are plausible, then (T), the theoretical structure of the individual demand curve analysis, can be applied to the aggregate demand curve analysis.

The main situations are the following, where the word "comprehensive" means "including all commodities and income," and the opposite of "household-level" is "aggregate":

1. The available data is comprehensive and household-level:

   It is appropriate to impose (T). There is no need to assume exact aggregation.

[12]The original result dealt with excess demand curves; it was later extended to market demand curves. See Wikipedia, *Sonnenschein-Mantel-Debreu theorem*, available at `https://en.wikipedia.org/wiki/Sonnenschein%E2%80%93Mantel%E2%80%93Debreu_theorem`.

Available demand systems: All the many demand systems surveyed by, for example, Barnett and Serletis.[13]

2. The available data is comprehensive but not household-level:

   (a) If the conditions for exact aggregation are not satisfied: it is inappropriate to impose (T).

   Available demand systems: As in Situation 1, but without the (T) restrictions.

   (b) If the conditions for exact aggregation are satisfied: it is appropriate to impose (T).

   Available demand systems: Only those consistent with exact aggregation are available. For exact linear aggregation, no plausible demand systems exist, as described in Section 3 below. For exact nonlinear aggregation, the demand systems are in the "generalized linear" or "GL" class, and include PIGL, PIGLOG, and AIDS, as described in Section 4 below.

3. The available data is not comprehensive and not household-level:

   (a) If the conditions for exact aggregation are not satisfied: it is inappropriate to impose (T).

   Available demand systems: As in Situation 1, without the (T) restrictions.

   (b) If the conditions for exact aggregation are satisfied: it is appropriate to impose some but not all parts of (T) if one has income data, and it is appropriate to impose no parts of (T) if one lacks income data, as described in Section 5.

   Available demand systems: As in Situation 2b.

From this overview, it is clear that a key question is: how plausible are the assumptions required for exact aggregation—in other words, how plausible are the assumptions required in order for the market demand curve to be thought of as being derived from "the average consumer" or from "one representative (but not mathematically average) consumer"? We answer this question in the next two sections, Section 3 for exact linear aggregation

[13] William A. Barnett and Apostolos Serletis, *Consumer Preferences and Demand Systems*, 147 JOURNAL OF ECONOMETRICS 210 (2008). Also William A. Barnett and Apostolos Serletis, *Measuring Consumer Preferences And Estimating Demand Systems* (2009). In Daniel J. Slottje, ed., QUANTIFYING CONSUMER PREFERENCES at Chapter 1, 1–35.

and Section 4 for exact nonlinear aggregation. Section 5 proves the assertions made about Situation 3. Econometric results that employ "standard" assumptions without specific justifications relevant to the actual antitrust case should be questioned.

3. Exact Linear Aggregation: The market demand curve can be thought of as being derived from "the average consumer"

There are two types of exact linear aggregation: "local," for which only small changes in income are contemplated, and "global," for which the aggregation must be able to be carried out for any level of income.[14] In Subsection A we ask what assumptions have to be made in order for us to be able to think of the market demand curve as being derived from "the average consumer," locally. In Subsection B we ask how realistic those assumptions are. Subsections C and D repeat Subsections A and B but for the global case. The conclusions we will come to are that the conditions for global exact linear aggregation are too narrow to be at all plausible, while the conditions for local exact linear aggregation are only somewhat more plausible. These dispiriting conclusions will prompt our interest in Section 4, where aggregation turns out to be possible under less restrictive, but still questionable, assumptions. An important conclusion is that antitrust econometrics can have hidden implausible assumptions. Because the mathematics are complicated, econometric results in antitrust cases may receive a lower level of scrutiny then they deserve. No judge would give a pass to an oral argument by a lawyer with a hidden implausible assumption. Empirical economists should be subject to equal scrutiny in the courtroom.

Throughout this Section we follow Section 6.1 of Deaton and Muellbauer[15] so closely that *we will identify text from that source merely by enclosing it in 'single' quote marks*. Our notation is slightly different from theirs, however.

A. Needed Assumptions for local exact linear aggregation.

Suppose there are H households, indexed by h, each of which has an income of I^h and a demand for good i of $d_i^h(I^h, \mathbf{p})$, where $\mathbf{p}$ is the list of prices $p_1, p_2, \ldots, p_n$ which all households face. (This notation, using $d_i^h(I^h, \mathbf{p})$ for the demand for good i by household h, with that demand depending on the household's income I^h and depending on all of the commodity prices $\mathbf{p}$, will

[14] Angus Deaton and John Muellbauer, Economics and Consumer Behavior (1980) at 150.

[15] *Id.*

be used extensively throughout this paper.) *Exact linear aggregation* is possible if and only if for every good i, average demand is some function—call it $\bar{d}$—of average income. This condition is expressed mathematically as:

$$\frac{d_i^1(I^1, \mathbf{p}) + d_i^2(I^2, \mathbf{p}) + \cdots + d_i^H(I^H, \mathbf{p})}{H} = \bar{d}\Big(\frac{I^1 + I^2 + \cdots + I^H}{H}, \mathbf{p}\Big)\,.$$

The problem for exact linear aggregation is that 'in general, no such function as $\bar{d}$ exists.' In order for any form of $\bar{d}$ to exist, note that $\bar{d}$ 'does not depend on the distribution of [incomes] I^h. Hence, for the equation to hold, a reallocation of a single unit of currency from any one to any other individual must leave market demands unchanged. This can only happen if [...] the marginal propensities to spend are identical for all consumers. Rich consumers must allocate changes in their outlay in exactly the same way as do poor consumers. This observation implies that the functions $d_i^h(I^h, \mathbf{p})$ must be linear in I^h, that is, for some functions α_i^h and β_i of $\mathbf{p}$ alone,

$$d_i^h(I^h, \mathbf{p}) = \alpha_i^h(\mathbf{p}) + \beta_i(\mathbf{p})\, I^h \tag{1}$$

where, although α_i^h is indexed by h, $\beta_i(\mathbf{p})$ is not.' (Equation (1) has the same form as the high school equation for a straight line, $y = mx + b$, with the left-hand side playing the role of y, $\alpha_i^h(\mathbf{p})$ playing the role of b, $\beta_i(\mathbf{p})$ playing the role of m, and I^h playing the role of x.)

Economists give the name "Engel curve" to the relationship between a household's income and its consumption of a good. When Engel curves are depicted in graphs, income is on the horizontal axis and demand (or "consumption") is on the vertical axis. On such a graph, when the Engel curve is upward-sloping, the good is "normal," that is, when income rises, its consumption rises. On such a graph, when the Engel curve is downward-sloping, the good is "inferior," that is, when income rises, its consumption falls. If income changes have no effect on a consumer's demand for a product, the Engel curve for the product will be a flat line.

Saying, as we did in the paragraph before last, that "$\beta_i(\mathbf{p})$ is not indexed by h" is equivalent to saying that for each good i, *the Engel curve's slope for commodity i is the same for every household*, and saying that β is not a function of I^h means that *the Engel curves are straight lines*. 'Suppose now that individuals maximize utility. In this case, (1) will hold if and only if' each consumer has quasi-homothetic preferences, which means that Engel

curves are perfectly straight lines. (Blundell and Stoker[16] attribute this result originally to a 1953 paper by Gorman.)

B. How realistic are the needed assumptions for local exact linear aggregation? The consensus is that these assumptions are quite unrealistic.

'Viewed as necessary conditions for aggregation, quasi-homothetic preferences, or equivalently, linear Engel curves, are extremely stringent. For example, any commodity not consumed at low budget levels is immediately excluded. Consequently, if linear aggregation is to work at all, it can only do so for broadly defined composites of goods.'

This is problematic in the antitrust context, where the goods are often so narrow, such as "different brands of dry cat food," that many households consume zero levels of many of them, even at high and moderate income levels, let alone at low income levels.

Barnett and Serletis[17] simply write, "Linearity in expenditure implies marginal budget shares that are independent of the level of expenditure, suggesting that poor and rich households spend the same fraction of an extra dollar on each good. This hypothesis, as well as the hypothesis of expenditure proportionality, are too restrictive for the analysis of household budget data."

We conclude that the assumptions needed for local exact linear aggregation are unrealistic, and economists ought not to assume local exact linear aggregation. (They usually do not.)

C. Needed assumptions for global exact linear aggregation. Thinking of (1) as an economic version of $y = mx + b$, we need both of the economic variables, income I^h (the analog of x) and quantity demanded $d_i^h(I^h, \mathbf{p})$ (the analog of y) to be positive or zero, not negative. That will require restrictions on the values of m and b, because unrestricted, $y = mx + b$ can certainly be negative, even if x is positive. This is the economists' next problem: to ensure that none of the d_i^h's in (1) can be negative. If, as sometimes happens, either $\alpha_i^h(\mathbf{p})$ or $\beta_i(\mathbf{p})$ is negative, 'the permitted range of I^h has to be restricted.' However, if 'we do not wish to place any restriction on the I^h's and' instead demand that 'aggregation be possible for all $I^h \geq 0$, we must delete the intercepts $\alpha_i^h(\mathbf{p})$ since otherwise some demands will be negative.' [Another line of reasoning: if $I^h = 0$ we should have $d_i^h = 0$, and this requires $\alpha_i^h \equiv 0$ if (1) holds.] 'Hence, this "global" aggregation implies

[16] Richard Blundell and Thomas M. Stoker, *Models of Aggregate Economic Relationships that Account for Heterogeneity* (2007). In James J. Heckman and Edward E. Leamer, eds., HANDBOOK OF ECONOMETRICS, VOLUME 6A at Chapter 68, 4619.

[17] *Supra* note 13 at 22–3.

that'

$$d_i^h(I^h, \mathbf{p}) = \beta_i(\mathbf{p})\, I^h\,. \tag{2}$$

This means that the Engel curves are all straight lines through the origin, and have the same slope for every household; budget shares $p_i d_i^h / I^h = p_i \beta_i(\mathbf{p})$ *are independent of total expenditure* (which Barnett and Serletis[18] point out contradicts "Engel's Law, according to which the budget share of food is smaller for rich than for poor households")*; and expenditure elasticities for all consumers and all goods are unity (meaning that every 1% increase in income induces every consumer to buy 1% more of every good).*[19] If this holds, it implies that

$$\frac{d_i^1(I^1, \mathbf{p}) + d_i^2(I^2, \mathbf{p}) + \cdots + d_i^H(I^H, \mathbf{p})}{H} = \beta_i(\mathbf{p}) \frac{I^1 + I^2 + \cdots + I^H}{H}\,,$$

so, indeed, average market demand (the left-hand side) is a function of average income ($(I^1 + I^2 + \cdots + I^H)/H$).

D. How realistic are the needed assumptions for global exact linear aggregation?

The restrictions needed for global exact linear aggregation are so wildly unrealistic that most economists shun (2) in favor of (1) and simply hope that their estimates of $\alpha_i^h(\mathbf{p})$ and $\beta_i(\mathbf{p})$, together with observed or realistic levels of I^h, never lead to a negative d_i^h.

4. Exact Nonlinear Aggregation: The market demand curve can be thought of as being derived from "a representative consumer"

Because Section 3's conditions for exact linear aggregation are so unrealistic, in this Section we investigate the less restrictive assumptions needed for exact nonlinear aggregation. This is an important topic because economists often impose theoretical restrictions from individual demand curves, restrictions we called (T) in Section 2, onto aggregation demand curves, and that is only justified if the conditions for exact linear or nonlinear aggregation hold; and given that we just concluded that the conditions for exact linear aggregation are very unlikely to hold, the conditions for exact nonlinear aggregation are quite important to know.

In Subsection A we ask what assumptions have to be made in order for us to be able to think of the market demand curve as being derived from "a representative consumer" (not by Section 3's "average consumer"). In

[18] *Supra* note 13 at 21.

[19] From Deaton and Muellbauer *supra* note 14 at 144, economists describe such preferences by saying that all the consumers' preferences are "homothetic."

Subsection B we ask how plausible those assumptions are. In Subsection C we ask what data is required to carry out econometrics based on those assumptions.

Overall, the conclusions are somewhat disappointing. While the conditions for exact nonlinear aggregation are certainly less restrictive than for exact linear aggregation, they are not trivial, and probably require commodities to be considered in large categories rather than analyzed individually. Also, it turns out that we will require household-level data in order to carry out exact nonlinear aggregation, but if we had household-level data, we probably would have little need to aggregate it. Therefore, the most likely situation will be one in which we wish to aggregate because we do not have household-level data, but lacking household-level data, we cannot carry out the aggregation exactly. Therefore, an aggregation bias will be present, meaning that the promise of exact aggregation (and thus justification for imposing restrictions (T)) will not be achievable in practice.

A. Needed assumptions for exact nonlinear aggregation. Throughout Section 4A we follow Section 6.2 of Deaton and Muellbauer so closely that *we will identify text from that source merely by enclosing it in 'single' quote marks*.

Exact nonlinear aggregation restricts average aggregate budget share for the ith good, $p_i(d_i^1 + d_i^2 + \cdots + d_i^H)/(I^1 + I^2 + \cdots + I^h)$, to depend on prices and, not on *average* income (because that leads to exact linear aggregation), but on a "representative" level of income which we will denote by I_0, 'which itself can be a function of the distribution of expenditures and of prices. If this holds, the market pattern of demand can be thought of as deriving from the behavior of a single representative individual endowed with income I_0 and facing prices **p**.'

In order for such a representative consumer to exist, one must place

> '... strong restrictions upon Engel curves; note, for example, that for each household the slopes of the different Engel curves will vary linearly with one another as total expenditure changes at constant prices. This does not, of course, imply that the Engel curves themselves are linear.
>
> 'Since these linear relations occur... the name given to the conditions for consistent nonlinear aggregation is generalized linearity (GL).... A particularly interesting special case occurs when the representative expenditure level is independent of prices and depends only on the distribution of expenditures. This case, known as price independent generalized linearity (PIGL) occurs

> when the microcost functions take the form [here follows an equation involving "α"...] When α tends towards zero, [that equation becomes the] form known as PIGLOG.'

Although Deaton and Muellbauer write that in order for a representative consumer to exist, strong restrictions have to be placed on Engel curves, they do not explain those restrictions beyond the first paragraph in the quotation displayed above. We now further explain those restrictions.

To begin, we have to translate into mathematical terms Deaton and Muellbauer's contention that 'for each household the slopes of the different Engel curves will vary linearly with one another as total expenditure changes at constant prices.' We also want to show that that condition is, as Deaton and Muellbauer claim, necessary and sufficient for existence of a representative consumer, because neither Deaton and Muellbauer nor anyone else seems to have published its proof. (All of the proofs for this paper are in the Appendix.)

Proposition 1. *There is a representative consumer if and only if there exist some numbers A_{hij} and B_{hij} such that*

$$\frac{\partial\, d_i^h(I^h, \mathbf{p})}{\partial I^h} = A_{hij}\, \frac{\partial\, d_j^h(I^h, \mathbf{p})}{\partial I^h} + B_{hij} \quad \text{for each } h \text{ and for all goods } i \neq j. \tag{3}$$

In words: the slope of the Engel curve for good i is equal to some constant "A_{hij}" times the slope of the Engel curve for good j, plus another constant "B_{hij}."

To help with interpretation, it is going to be useful to have two more results:

Proposition 2. *Suppose a representative household exists. Then the ratio of the second derivative of household h's Engel curve for good i to the second derivative of household h's Engel curve for good j is a constant.*

Furthermore, if a representative household exists and if, for some household h, there exists an interval of incomes $\hat{I}$ on which d_j^h is a linear function of I (for fixed $\mathbf{p}$), then for that household and for all other goods $i \neq j$, d_i^h is a linear function of I on $\hat{I}$.

Corollary. *If for some household h, there exists an interval of incomes $\hat{I}$ on which d_j^h is a constant function of I (meaning that on that interval, household h's consumption of good j stays the same as household h's income changes), then for that household and for all other goods $i \neq j$, d_i^h is a linear function of I on $\hat{I}$.*

Proposition 2's second sentence implies that *if both of a pair of Engel curves are concave (or convex) for one value of I^h, they are concave (or convex) for all values of I^h*. Similarly, *if one of a pair of Engel curves is concave and the other is convex for one value of I^h, then they will have opposite convexity for all values of I^h*.

B. How plausible are the restrictions needed for exact nonlinear aggregation? The convexity or concavity restrictions just described seem completely arbitrary. There is no reason to assume consumers behave in that way.

The above Corollary is key to deeper understanding of how restrictive Proposition 1 is. Suppose good j is Iams cat food and suppose that when household h's income I^h is greater than \$100,000, the household's demand for Iams cat food is constant. Then the Corollary tells us that the Engel curves for *all* other goods are linear when $I^h > \$100,000$. This is uncomfortably close to the "Engel curves are linear in I for all values of I^h" assumption which we rejected as being unrealistic when discussing exact linear aggregation. However, here, it is only household h's Engel curves which are linear, and only for $I^h > \$100,000$; for exact linear aggregation, all households' Engel curves had to be linear, for all income levels.

For an even more stark example, suppose good j is Iams cat food and suppose household h has no cats. Then for all levels of I^h, this household's demand for Iams cat food is constant (at zero). Reasoning as before, this means that the Engel curves for all other goods are linear for *all* values of I^h. This *is* the "Engel curves are linear in I for all values of I^h" assumption which we rejected as being unrealistic when discussing exact linear aggregation, although again, here it is only household h which is unrealistic, not all the households; nevertheless, household h *is* unrealistic in this situation.

To summarize, if there is a commodity which a household never consumes, then all the Engel curves of that household are linear for all values of income. If, as income rises, the consumption of one good becomes constant, then the Engel curves of that household for all the other goods from that income level on have to be linear. It follows that if, above a certain level of income, the consumption of one good becomes constant for all households, then the Engel curves of all households for all the other goods from that income level on have to be linear. For example, if above an income level of \$500,000, the demand for Iams cat food is constant for every household, then the Engel curves of all households for all other goods have to be linear for incomes above \$500,000. As Deaton and Muellbauer say: "strong restrictions." Indeed, econometricians almost never even study whether or not Engel curves are linear, instead studying the different question of whether

or not they are log-linear (that is, whether the logarithm of consumption is a linear function of the logarithm of income)—a question that is irrelevant to deciding whether aggregation is appropriate or not.

If one wants to avoid forcing all Engel curves to be linear over a range of income levels, one needs to avoid having any commodity's consumption be flat (or linear) on that range of income levels. Avoiding flat consumption would be quite challenging for goods that are inferior at low or moderate levels of income: many of them probably stay at zero consumption for all large incomes (or stay constant at some small level there). It is probably necessary to define "goods" i, j, ..., quite broadly in order to ensure that no good's Engel curve has a flat (or linear) portion.

It is true that having all Engel curves be linear over a *small* range of incomes is probably unobjectionable. However, outside of defining the goods broadly, it is unclear how one would ensure that such a range of incomes would be small. Moreover, even if linearity does not occur, one still has Proposition 2's counterintuitively linked concavity and convexity conditions standing in the way of accepting exact nonlinear aggregation as being plausible.

C. What data is needed to implement exact nonlinear aggregation? As mentioned above, a representative household exists if and only if preferences belong to the "generalized linearity" (GL) class, and special cases within the GL class are the PIGL class and the PIGLOG class. Within the PIGLOG class is the "Almost Ideal Demand System" (AIDS), which is the most commonly used econometric approach to market demand estimation in antitrust; see Deaton and Muellbauer.[20]

Since the AIDS class of preferences is a subset of the GL class, if one assumes AIDS preferences, then aggregate preferences can be generated by a representative consumer, and there will be no aggregation bias. This makes the title of a 1996 paper by Mittelhammer, Shi, and Wahl initially puzzling: "Accounting for Aggregation Bias in Almost Ideal Demand Systems." The problem turns out to be that in order to calculate the AIDS system, you need to know not only that a representative consumer exists, you also have to be able to calculate who the representative consumer is, and that requires household-level data which, if you had it, would put you in Section 2's Situation 1, where there is no need to use AIDS nor to have a representative consumer. Here is the way this is explained by the authors of the above

[20] Angus Deaton and John Muellbauer, *An Almost Ideal Demand System*, 70/3 THE AMERICAN ECONOMIC REVIEW 312 (1980) at 313.

1996 paper in a 1994 working paper with almost the same title which lists the authors in a different order.[21]

> Clearly the calculation [for the AIDS model] of the weighted geometric mean of expenditures, x^*, in the aggregate share equation *requires detailed information on the distribution of total expenditures over consumers*. Unfortunately, most empirical data available for applied demand studies are measured at an aggregated level, and the information necessary for computing x^* is often not available in practice. *In demand studies utilizing aggregate time series data, researchers often use the simple average of individual expenditures (i.e., per capita expenditure) to replace the geometric mean*. Deaton and Muellbauer (1980a and 1980[b]) have shown that if the *average* aggregate budget share is to be specified as a function of prices and per capita expenditure, this requires the restrictive conditions of exact linear aggregation. In the case of exact (price-independent) nonlinear aggregation, such as AIDS, it is required that the aggregate budget share, $\bar{w}_i$ depend on prices and a representative level of total expenditure x_0 which itself depends on the distribution of expenditures. In this case, "the market pattern of demand can be thought of as deriving from the behavior of a single representative individual endowed with total expenditure x_0 and facing prices p." (Deaton and Muellbauer 1980a p. 154). In the case of PIGLOG preferences, it is clear from (6) that the appropriate level of representative expenditure is given by $x_0 = \prod x_h^{r_h}$. It follows that *using $\bar{x}$ [the average value of expenditures x] in place of x_0 constitutes a misspecification of the AIDS model*.[22]

This confirms that the AIDS model itself has no aggregation bias, but if one misspecifies it, then the misspecified AIDS model has an aggregation bias. The obvious solution—do not misspecify the AIDS model—fails whenever one *lacks the household-level data required to correctly specify the AIDS model*. The authors' 1996 paper suggests an approximate solution: supplement one's aggregate data set with a different, household-level data set.

---

[21] Thomas I. Wahl, Ron C. Mittelhammer, and Hongqi Shi, *Accounting for Aggregation Bias in Empirical Demand Models—The Case of Almost Ideal Demand Systems* (1994). Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, Illinois, at 194.

[22] Compare the "x" in (16) of Deaton and Muellbauer *supra* note 20 with the "$\bar{x}$" in (20) of that paper (it at 314 says that "$\bar{x}$ is the average level of total expenditure x_h").

> . . . it is evident that, to calculate the expenditure aggregation bias term, time-series information on the number of households and on individual households' shares of aggregate expenditure are needed. Information on the shares of aggregate expenditure across households is generally unavailable or inaccessible. However, time-series information on the number of households in different income categories is readily available for most developed economies and can provide valuable information for closely approximating the income distribution and aggregation bias term in the aggregate AIDS model. (p. 250)

To summarize: to use the AIDS model, one needs to have household-level data, in order to calculate the "representative expenditure" (otherwise you are misspecifying the AIDS model and there will be aggregation bias). But referring back to the three data availability Situations given in Section 2, if one had household-level data, one would be in Situation 1, in which one would not need to use AIDS, nor be bothered about existence of a representative consumer, nor about aggregation at all. Many economists want to use AIDS, and assume a representative consumer exists, when they do not have household-level data (Situations 2b and 3b of Section 2); but we now learn that one cannot use AIDS (at least not exactly) in Situations 2b and 3b because one cannot calculate the "representative expenditure." The three co-authors suggest an approximation to get out of this Catch-22: substitute "representative expenditure" calculated from economy-wide income distribution data, which is widely available, in the place of the "representative expenditure" for the households that generated one's own data. Essentially, this is a work-around which enables one to proceed in Situations 2b and 3b. More recent literature uses "aggregation factors" to describe "the degree of bias in recovering (individual) price and income elasticities from aggregate data alone," as Blundell and Stoker[23] put it.

If lack of household-level data causes one to be unable to exactly implement AIDS estimation, the other response is to abandon Situations 2b and 3b and turn instead to 2a and 3a. That is, abandon assuming a representative agent, thus abandon any justification for imposing (T). Stoker[24] calls this abandoning "descriptions that are straightjacketed by the capricious enforcement of restrictions of optimizing behavior by a single individual." Freed from the straightjacket imposed by wanting to use (T), one can then

[23] *Supra* note 16, at 4621.

[24] Thomas M. Stoker, *Empirical Approaches to the Problem of Aggregation Over Individuals*, 31/December JOURNAL OF ECONOMIC LITERATURE 1827 (1993) at 1829.

use any of the demand systems described in Situations 2a and 2b, and so one can use a demand system which, unlike AIDS, "seem[s] to do a good job of fitting the data, such as the QUAIDS system of Banks, Blundell and Lewbel (1997)" (*supra* note 16 at 4622). On the other hand, the straightjacket was extremely helpful in shrinking down from $(n+2)(n-1)$ the number of parameters that needed to be estimated, so the straightjacket will be missed in some situations.

Nevertheless, Blundell and Stoker's outlook for econometricians who do not have any household-level data (Situations 2 and 3) is grim:

> While we have advanced the idea of using aggregation factors (derived from time-series of individual data) to summarize the impacts of aggregation, the specific method one uses is less important than the ability to use all available types of information to study economic relationships. That is, it is important to study any relationship among economic aggregates with individual data as well as aggregate data, to get as complete a picture as possible of the underlying structure. Even though modeling assumptions will always be necessary to develop explicit formulations of aggregate relationships, testing those assumptions is extremely important, and is not possible without extensive individual data over sequential time periods.[25]

Stoker[26] explained the basic problem a few years earlier:

> Models that account for individual heterogeneity will typically not be estimable using data on economy-wide averages alone; additional data on distributional composition. . . , or micro data on individual behavior, will need to be incorporated. This should come as no surprise; to study relations that involve heterogeneous individual responses without distributional information is analogous to studying dynamic relations without using data over time. [at 1836] [. . .]
>
> Whether a representative agent model fits the data or not, there is no realistic paradigm where the parameters of such a model reflect only behavioral effects, uncontaminated by compositional considerations. The application of restrictions appropriate for individual behavior directly to aggregate data [that is, applying (T)] is a practice without any foundation, and leads to

[25] *Supra* note 16 at 4658.

[26] *Supra* note 24.

> biases that are impossible to trace or measure with aggregate data alone. [at 1870]
>
> Hand-in-hand with the necessity of using all relevant data is the necessity of checking or testing all relevant assumptions underlying a model. Aside from a platitude of good empirical work, it is important to stress the testing aspect here because altogether too little attention has been paid to checking or testing assumptions required for aggregation, relative to assumptions on the form of individual behavioral models. [at 1870] [. . .] Approaches that neglect individual heterogeneity, such as pure representative agent modeling, should be abandoned. [at 1871]

One cannot test the assumptions required for a representative consumer using only aggregate data. When one tests them using household-level data, what does one find? Barbett and Serletis say, rather mildly, "most of the commonly used PIGLOG specifications are of rank two, and thus do not have enough flexibility in modelling the curvature of Engel curves with large variations in income" (2008 at 218). Banks, Blundell, and Lewbel[27], writing after extensive use of household-level data from the U.K. Family Expenditure Survey, say that their "quadratic logarithmic class nests both the Almost Ideal (AI) model of Deaton and Muellbauer and [. . .]. Unlike these demand models, however, the quadratic logarithmic model permits goods to be luxuries at some income levels and necessities at others. The empirical analysis we report suggests that this is an important feature. . . . The specific form we propose—the Quadratic Almost Ideal Demand System (QUAIDS)—is constructed so as to nest the AI model and have leading terms that are linear in log income while including the *empirically necessary* [emphasis added] rank 3 quadratic term" (which AIDS lacks). Stoker[28] is the most emphatic; he says about AIDS (emphasis mine):

> In particular, (4.12) [an equation "used in Deaton and Muellbauer's (1980[b]) estimation"] rests on the assumption that a) (4.7) [the AIDS demand system] is valid, with no individual heterogeneity in demands aside from income effects and b) [. . .]. Each of these assumptions is testable *with micro data* [emphasis added], and patently unrealistic. . . .

[27] James Banks, Richard Blundell, and Arthur Lewbel, *Quadratic Engel Curves And Consumer Demand*, 79/4 THE REVIEW OF ECONOMICS AND STATISTICS 527 (1997) at 527–8.

[28] *Supra* note 24 at 1855.

Overall, then, it is not only the representative agent approaches 2b and 3b which have sustained heavy criticism, but *any* attempt to proceed in Situations 2 and 3. Yet empirical antitrust work typically deals with Situation 3, to which we now turn.

5. Demand Estimation when Lacking Data on Some Goods

Ulrick[29] describes the sort of data typically available in antitrust analysis:

> The AIDS model above is often estimated with Nielsen and IRI scanner data to generate brand-level demand elasticities. Scanner data are particularly suited for this type of analysis. The data are almost always available weekly by SKU and city. The data will include total retail dollars, equivalised units, units, and marketing variables. The data generally cover food stores (the grocery channel), but in some cases mass merchandiser data is also available (ie Target, Kmart).

This is similar to the data available to Hausman, Leonard, and Zona.[30] In Section 2, we referred to these settings as Situation 3.

The question Situation 3 raises is to what extent one can impose the restrictions (T)—remember many econometricians *do* impose (T)—if one lacks data on many goods the household purchases, and lacks data on the household income. To answer this, we need to go one-by-one through the list of restrictions which constitute (T). Up to this point, we have avoided saying what the restriction list (T) actually contains, but now, not only do we have to list the restrictions, but we have to mathematically check each one to make sure they still apply if one only has data on some commodities. What we will find is that two of the restrictions in (T) fail to be applicable to the case when one lacks data on some goods; the other two restrictions in (T) do apply to that case, but only if income data is available, and if income data is not available, those restrictions should not be imposed either. This calls into question the validity of many econometric studies in which (T) is imposed.

We follow the order of restrictions in Ulrick's treatment.[31]

Adding up. This is the restriction that the sum of expenditures, $p_1 d_1^h + p_2 d_2^h + \cdots + p_n d_n^h$, is equal to household h's income; or the corresponding aggregate restriction that the sum of expenditures of all the households is

[29] *Supra* note 6 at 138–9.

[30] *Supra* note 7.

[31] *Supra* note 6, Section C.

equal to the sum of all the household's incomes. If one only has data on, for example, commodities 1 and 2, there is no restriction to impose, because expenditures on commodities 1 and 2 are not constrained to be any particular number. There is one potential work-around. Suppose commodities 1 and 2 are two types of pet food, suppose none of the other commodities are pet foods, and suppose that regardless of how much prices and income change, this household always spends a fixed amount of money on pet food. Then $p_1 d_1^h + p_2 d_2^h$ would always equal this fixed amount of money, and this would function like the more general adding up restriction. The problem is that the assumption "regardless of how much prices and income change, this household always spends a fixed amount of money on pet food" is probably incorrect.

It is true that in the context of multi-stage budgeting (as in Hausman, Leonard, and Zona[32]), assuming a fixed expenditure on, say, pet food may not be problematic. Multi-stage budgeting implies that preferences are not of the AIDS or GL form (but instead have some separability properties[33]). Hausman, Leonard, and Zona say "our econometric specification at the lowest level is the 'almost ideal demand system' of Deaton and Muellbauer,"[34] but to be clear, this only means that it uses the AIDS form for one of their three budgeting stages; their consumers do not, overall, have AIDS preferences, so no representative consumer exists and they should not impose (T). (Unfortunately they do impose one component of (T), symmetry (see pages 163 and notes on Tables 2–4).) On the other hand, the true adding up condition itself (ignoring the other elements of (T)) just comes from the budget constraint, and it applies in the most general cases: it is one of the few conditions which the market demand curve has even in the Sonnenschein-Mantel-Debreu theorem.

Note that even the lowest-level model of Hausman, Leonard, and Zona is not a true AIDS model because its dependent variable is, for example, the amount of money spent on Budweiser premium beer as a fraction of the amount of money spent on all premium beers, whereas in a true AIDS model, the corresponding dependent variable would be the amount of money spent on Budweiser premium beer as a fraction of the amount of money spent on all commodities. With data like Hausman, Leonard, and Zona's, and no additional data on incomes, there is no way to know how much money the

[32] *Supra* note 7.

[33] See Deaton and Muellbauer, *supra* note 14, Chapter 5.

[34] *Supra* note 7 at 162.

consumers are spending on all commodities, so there is no way to construct a true AIDS model, so there is no justification for (T).

Homogeneity. This restriction requires that quantity demanded be unchanged if one multiplies all prices and income by a constant. If household consumes more than (for example) two commodities but one only has data on the first two commodities, one cannot impose this restriction. After all, if the household consumes more than two commodities then it is false that "if only the prices p_1 and p_2 are multiplied by a constant, d_1^h and d_2^h are unchanged."

Symmetry. The Symmetry condition is one of the most commonly-imposed parts of (T), but unfortunately, there is no intuitive, non-mathematical interpretation of it. The Symmetry condition is that, for all goods i and $j \neq i$,[35]

$$\frac{\partial d_i^h(I,\mathbf{p})}{\partial p_j} + \frac{\partial d_i^h(I,\mathbf{p})}{\partial I} d_j^h(I,\mathbf{p}) = \frac{\partial d_j^h(I,\mathbf{p})}{\partial p_i} + \frac{\partial d_j^h(I,\mathbf{p})}{\partial I} d_i^h(j,\mathbf{p}) \qquad (4)$$

or, at the aggregate level, (4) dropping the household indexes h. Notice that the only difference between the two sides of (4) is that the i and the j are interchanged, which gives the condition its name, symmetry. Notice also that if Engel curves were flat, the second terms on each side of (4) (called the "income effect" terms) would be zero, and (4) would collapse to $\partial d_i^h/\partial p_j = \partial d_j^h/\partial p_i$, that is, "the increase in purchases of good i when the price of good j changes is equal to the increase in purchases of good j when the price of good i changes," which is as close to an intuitive explanation of (4) as one is likely to get. (The Appendix shows how to express the symmetry condition in more conventional but more indirect way; see its Proposition 3.)

The Symmetry condition should hold even if one has data only on a few of the commodities which the consumer purchases. However, the income effect ($\partial d/\partial I$) terms in (4) could only be calculated if one had data on income (at the household or, assuming a representative agent, the aggregate level). If data on income is lacking, as is often the case, then even though the restriction (4) should hold, there would be no way to do the calculations necessary to impose it.

If the Symmetry condition ought to hold, it is a considerable help: Muellbauer[36] writes, "Of the restrictions implied by utility theory, by far the most

[35] Ulrick and most other authors express this using "Hicksian," or "compensated," demand functions, instead of the "Marshallian" demand functions used in this paper. The expression using Hicksian demand functions is much more compact.

[36] *Supra* note 8 at 525.

important saving in parameters results in the $\frac{1}{2}n\,(n-1)$ restrictions implied by the symmetry of compensated cross-price effects," which is this condition. This is why it is imposed so often, even though its imposition is apparently not always theoretically justified.

Negativity. Negativity has two aspects. The first is that for all goods i,

$$\frac{\partial d_i^h(I,\mathbf{p})}{\partial p_i} + \frac{\partial d_i^h(I,\mathbf{p})}{\partial I} d_i^h(I,\mathbf{p}) \leq 0\,. \tag{5}$$

In the absence of income effects, this simply says that "demand curves are downward-sloping." In the presence of income effects, it is difficult to fruitfully express (5) in words. This aspect of the Negativity restriction should hold even if one has data only on a few of the commodities which the consumer purchases, although just like for the almost identical terms in (4), income data is needed to compute the left-hand side.

The second aspect of Negativity involves more mathematics, and the reader unfamiliar with quadratic forms is invited to skip this paragraph. Let the left-hand side of (4) be abbreviated S_{ij}. If one lacks data on how demand varies when income varies, the S_{ij}'s cannot be calculated and "Negativity's second aspect" restrictions cannot be imposed. Otherwise, with n being the number of commodities, denote by $\mathbf{S}$ (sometimes called the "Slutsky Matrix" or the "Substitution Matrix") the $n\times n$ matrix whose (i,j) element is S_{ij}. The restriction is that $\mathbf{S}$ be negative semidefinite (and symmetric, which underlies the above symmetry restriction). It can be shown that the submatrix of $\mathbf{S}$ obtained by retaining only some of its rows and the corresponding columns should also be negative semidefinite, so this restriction does carry through to the case where there is data only on some commodities.[37]

6. Consumer Surplus and Demand-curve Estimation

One purpose for which econometric estimates of demand and supply curves in antitrust are performed is to make judgments about economic welfare changes. If this is done using Kaldor's Compensating Variation and Hicks's Equivalent Variation, no restrictive assumptions about consumer preferences have to be made. However, if this is done instead using consumer surplus

[37] Proof: Reorder the commodities so that the commodities one has data on are the first ones. Suppose there are $n' < n$ such commodities. Since $\mathbf{S}$ is negative semidefinite, all of its principal minors of order r alternate in sign beginning with ≤ 0 for $r = 1, 2, \ldots, n$. This means that all of $\mathbf{S}$'s principal minors of order r alternate in sign beginning with ≤ 0 for $r = 1, 2, \ldots, n' < n$. The latter means that the submatrix of $\mathbf{S}$ obtained by retaining only n' of its first few rows and the corresponding columns is also negative semidefinite.

as the welfare measure, it has been known for a long time—in some sense, since Alfred Marshall[38]—that the commodity being studied must have no income effect, that is, that its consumption must not vary with income and therefore that its Engel curve must be horizontal:

$$d_i^h(I^h, \mathbf{p}) = \alpha_i^h(\mathbf{p})\,. \tag{6}$$

It is helpful to contrast this with the other assumptions discussed in this paper. For exact global aggregation, we needed

$$d_i^h(I^h, \mathbf{p}) = \beta_i(\mathbf{p})\, I^h \tag{2}$$

"Engel curves are all straight lines through the origin, and have the same slope for every household." For exact linear aggregation, we needed

$$d_i^h(I^h, \mathbf{p}) = \alpha_i^h(\mathbf{p}) + \beta_i(\mathbf{p})\, I^h\,, \tag{1}$$

"the Engel curve's slope for commodity i is the same for every household, and Engel curves are straight lines." And for exact nonlinear aggregation, we needed

$$\frac{\partial}{\partial I^h}\, d_i^h(I^h, \mathbf{p}) = A_{hij}\, \frac{\partial}{\partial I^h}\, d_j^h(I^h, \mathbf{p}) + B_{hij} \tag{3}$$

for each h and for all goods $i \neq j$: for each household, the slopes of the household's Engel curves for different commodities will vary linearly with one another as total expenditure changes at constant prices (and the implications of this in Proposition 2 and its corollary). While (6) clearly contradicts (2), at first glance it seems to be compatible with (1), setting $\beta_i(\mathbf{p}) = 0$, and it seems to be compatible with (3), which it turns into $0 = 0 + B_{hij}$, setting $B_{hij} = 0$. However, (1) has to hold for every commodity i, and (3) has to hold for every pair of commodities i and j, whereas (6) cannot hold for every commodity i, because then income levels would affect no one's consumption for any commodity, which is completely at odds with empirical findings. One could have (1) hold for all commodities and have $\beta_i(\mathbf{p})$ be zero for a subset of commodities, so that consumer surplus would be an exact welfare measure on that subset. One could also have (3) hold for all commodities and have B_{hij} be zero for a subset of commodities on which

[38]Marshall wrote that "In regard to different people allowance may have to be made where necessary for differences of sensibility and for differences of wealth: but it is seldom needed in considering large groups of people" in his margin notes for pages 130 and 131, Book III ("On Wants and their Satisfaction") Chapter VI ("Value and Utility") Section 3, of ALFRED MARSHALL, PRINCIPLES OF ECONOMICS, 8TH EDITION (1920).

(6) also holds, so that consumer surplus would be an exact welfare measure on that subset. Before taking these steps, though, one should first determine whether (6) is at all a plausible description of the way consumers buy the good i of interest. If i is an important good, as is likely if it is of antitrust interest, then the plausibility of (6), with consumption independent of income, may be in doubt.

7. "Do Firms Maximize Profit?" is an Empirical Question

Consumer behavior has been studied by empirical economists in many different subdisciplines for many different reasons, so remaining controversies tend to involve mathematical issues of the sort treated in this paper so far. The only non-mathematical, more conceptual controversies in consumer behavior revolve around the important question of whether the neoclassical assumption that consumers are rational is accurate, and if it is not, in what ways consumers fail to be rational. Behavioral economists study this question, but it is beyond the scope of this paper, beyond noting that if the commodities being studied are uncomplicated ones with which consumers are quite familiar, such as beer or pet food or breakfast cereal, the rationality assumption is a perfectly acceptable working hypothesis.

Antitrust analysis is not limited to demand behavior. Firms under antitrust scrutiny often contend that the conduct under review has an efficiency motive. Therefore, to judge the social consequences of an antitrust action we need to understand not only how consumers will be affected by it but how firms will be affected by it, and this requires understanding firms' behavior, which in turn is determined by firms' motivations. The neoclassical assumption is that firms are motivated by only one thing: profit[39] (or, in multi-period settings, "the net present value of profit," discussion of which is postponed until Subsection C). There are reasons to think that this answer is wrong. This Section examines those reasons.

In principle, firms could have both pecuniary and non-pecuniary motivations. Subsection A deals with firms' non-pecuniary motivations, arguing that many firms have such motivations, contrary to the economists' presumption, and that many business-people believe they should have such motivations. Subsection B deals with firms' pecuniary motivation, arguing that economists are completely correct that firms' pecuniary motivation *should be* profit, but that economists are quite incorrect in assuming that the pecuniary motivation of large firms in the U.S. in the present era *is* profit.

[39] DENNIS W. CARLTON AND JEFFREY M. PERLOFF, INDUSTRIAL ORGANIZATION, SECOND EDITION (1994) at 15–16. Obviously this does not pertain to non-profit firms.

Subsection C does for multi-period settings what Subsection B does for single-period settings.

A. Firms' non-pecuniary motivations. As noted above, most economists assume that firms' only objective is to maximize profit. This is the positive version of the normative "Friedman Doctrine," which holds that firms' only objective ought to be to maximize profit, or what in business is now called "shareholder value."[40] The normative version is true under the Arrow-Debreu assumptions of general equilibrium theory, assumptions which in no way describe our actual economy; unfortunately, economists have written very little on the natural follow-up question, namely what firms *ought to do* in our actual economy. There is no theorem asserting that in our actual economy firms ought to maximize profit, and, on the contrary, the entire field of Industrial Organization economics largely exists because in an economy which is not perfectly competitive, profit-maximizing firms do not behave in socially optimal ways.

The positive version of the Friedman Doctrine was not universally true in Milton Friedman's time—otherwise he would not have complained that it was not being followed—nor is it universally true today, since some firms both in his time and in ours act with broader social goals in mind. For example, a portrait of Merck & Co.'s George Merck was on the cover of Time magazine on August 18, 1952, together with his quotation, "Medicine is for people, not for profits."[41] It was this background that Friedman rebelled

[40] Milton Friedman, *A Friedman doctrine—The Social Responsibility Of Business Is to Increase Its Profits*, NEW YORK TIMES MAGAZINE (1970): "There is one and only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game." See also Wikipedia, *Friedman doctrine*, available at `en.wikipedia.org/wiki/Friedman_doctrine` and Thomas Carson, *Friedman's Theory of Corporate Social Responsibility*, 12/1 BUSINESS & PROFESSIONAL ETHICS JOURNAL 3 (1993)

[41] See *Time Magazine Cover: George W. Merck*, August 18, 1953, available at `web.archive.org/web/20061207070924/http://www.time.com/time/covers/0%2C16641%2C19520818%2C00.html`. For the full quotation see `www.merck.com/company-overview/history`, for the year 1950:

> In a defining moment for the company, George W. Merck gave a talk at the Medical College of Virginia at Richmond, during which he made a famous statement about how the medical and pharmaceutical community could be successful:
>
> "We try to remember that medicine is for the patient. We try never to forget that medicine is for the people. It is not for the profits. The profits follow, and if we have remembered that, they have never failed to appear."

against. As recounted by Scott Tong of marketplace.org:[42]

> Friedman and his University of Chicago free-market colleagues argued that corporations were taking on too many "social responsibilities": providing jobs, helping to fight pollution and reducing discrimination in society. In their eyes, the model was inefficient—and unfair to shareholders.
>
> "We saw enormous amounts of waste going on," said economist Michael Jensen, a former University of Chicago student of Friedman's now retired in Sarasota, Florida. Jensen went on to co-author the often-cited business paper. . . "Theory of the Firm," [in which] he and co-author William Meckling argued that corporate shareholders were shortchanged by corporate managers seeking perks. . . .
>
> Jensen went on to become a superstar professor at Harvard Business School where his ideas went mainstream. . . .
>
> So into the business playbook went three words: maximize shareholder value. Jensen may not have used those words, but to many that's how his ideas were applied by a entire generation of business leaders. . . .
>
> "There is a widespread and completely erroneous belief out there that there is some sort of legal duty that corporate managers have to 'maximize profits' or 'maximize shareholder value,' " said Cornell law professor Lynn Stout, author of "The Shareholder Value Myth". . . . "You can just pick up the Supreme Court case 'Hobby Lobby' decided just a few years ago," she said. "Read the majority opinion, where Justice Alito says, and I quote, 'modern corporate law does not require for-profit corporations to pursue profit at the expense of everything else.' ". . .
>
> Which brings us back to the man widely acclaimed as the intellectual father: Mike Jensen.
>
> "Has it happened the way I wanted it to happen? Eh, probably not," Jensen said. "There's always going to be some people who take it too far. And then cause damage."
>
> Jensen said focusing solely on stocks and stockholders is a "misreading" of his scholarship. . . . "I wouldn't put shareholders at the center," he said. "I'm still unhappy about the situation

[42] Scott Tong, *This is how Shareholders got to be First in Line for Profits*, BUSINESS INSIDER (2016) at https://www.businessinsider.com/the-story-of-shareholder-value-2016-6.

> where people end up thinking that shareholders are primary. That they are our only bosses. No."

Jensen is claiming that he was disturbed by managers' spending on themselves, but, unlike Friedman, not by managers' spending on "social responsibilities."[43]

Kangas[44] cites the strong condemnation of the Friedman Doctrine in 1971 by the "The Committee for Economic Development," which is part of the large-business organization called The Conference Board.[45] Almost 50 years later, condemnations of the Friedman Doctrine have come not only from the Columbia [University] Center on Sustainable Development, [46] but also from an author with the Booth School of Business at the University of Chicago[47] and another author from the Sloan School of Management at MIT.[48] All these sources give evidence of some firms acting in ways motivated not by profit but by improving society in general, by such actions as voluntarily reducing pollution, or setting up affirmative action programs, or making important charitable donations, or speaking out against laws disadvantaging racial minorities. In 2016, *Business Insider* echoed Prof. Stout's earlier language, publishing an article calling shareholder value a "myth."[49]

It follows that instead of simply assuming firms have only pecuniary motivations, antitrust economists should undertake empirical investigation

[43]Carson *supra* note 40 claims Friedman himself had two different notions of shareholder value, rather than one consistent notion.

[44]Alexander A. Kangas, *The Collective Action Problem of Capitalists and the Relative Autonomy of the State* (2015). Ph.D. dissertation, Department of Economics, University of Utah.

[45]COMMITTEE FOR ECONOMIC DEVELOPMENT, SOCIAL RESPONSIBILITIES OF BUSINESS CORPORATIONS (1971). Note that `www.ced.org/reports/social-responsibilities-of-business-corporations` says "CED Trustees are chief executive officers and key executives of leading US companies."

[46]Lisa Sachs, Jeffrey Sachs and Nathan Lobel, *Corporations Need to Look beyond Profits* (2017). Available at `academiccommons.columbia.edu/doi/10.7916/D8D79PV2`.

[47]Amy Merrick, *Is the Friedman Doctrine still Relevant in the 21st century? A year of crises has heightened the debate about what corporations owe society* (2021). Available at `https://review.chicagobooth.edu/economics/2021/article/friedman-doctrine-still-relevant-21st-century`.

[48]Neal A. Hartman, *Social Responsibility Matters to Business—A different view from Milton Friedman from 50 years ago* (2021). Available at `https://mitsloan.mit.edu/experts/social-responsibility-matters-to-business-a-different-view-milton-friedman-50-years-ago`.

[49]Dan Bobkoff, *Business Schools Led the Way in Promoting a Commonly Held Myth—and Now They're Trying to Undo That*, BUSINESS INSIDER (2016), available at `https://www.businessinsider.com/business-schools-led-the-way-in-promoting-shareholder-value-myth-2016-6`.

to determine whether or not this is so, for the particular firms in question. If it is not so, then that firm's other motivations and activities should be taken into consideration before allowing such a management team to be displaced by another which *is* solely driven by pecuniary concerns. Otherwise, the result can be shareholder-driven mergers whose aim is to increase firm cash flow and its distribution—misleadingly labelled "efficiencies"—but whose effect is to decrease socially beneficial activities.[50]

B. Firms' pecuniary motivations. Standard neoclassical economics goes on the assumption that firms' pecuniary motivation is profit,[51] but, especially as concerns large U.S. firms in the current era, this should not simply be accepted without empirical verification. Neoclassical economists' support for their position is weak: it is little more than a vaguely Darwinian argument that if some firms did maximize profit and others did not, then in the long run, the ones which maximized profit would out-compete the others, taking them over or running them out of business.[52] This argument is weak because many other things happen in the long run to muddy the situation. One of these confounding factors is the ongoing death of the experienced mortals who run businesses and their replacement with inexperienced ones, bringing a constant flow of new actors into the arena and raising the question of whether there will be any firms which *always* maximize profit (education not being perfect, as any teacher knows). Another confounding factor is novel technological innovations, each of whose profit-maximizing utilization may not be understood for decades. Another confounding factor is the likely inability of investors to tell when a firm which is not maximizing short-run profit actually *is* maximizing a longer-run version of profit (Subsection C's "net present value of profit"). Yet another reason to doubt that firms profit-maximize is the very mixed evidence of the effect of mergers on firms.[53]

---

[50]See discussion in WILLIAM LAZONICK AND JANG-SUP SHIN, PREDATORY VALUE EXTRACTION: HOW THE LOOTING OF THE BUSINESS CORPORATION BECAME THE US NORM AND HOW SUSTAINABLE PROSPERITY CAN BE RESTORED (2020).

[51]*Supra* note 39.

[52]*Supra* note 39 at 16.

[53]See Robert H. Lande and Sandeep Vaheesan, *Preventing the Curse of Bigness Through Conglomerate Merger Legislation*, 52 ARIZONA STATE LAW JOURNAL 75 (2020) at 102–104 notes 131–134. Besides giving an overview of the effect of mergers on firms, they write: "For an older survey see Dennis W. Carlton and Jeffrey M. Perloff, Modern Industrial Organization (4th ed. 2004) ('Additional research on profits subsequent to consolidation, not on stock price, is needed to confirm these efficiency gains. Without such research, some may argue that mergers and takeovers create illusory stock market value that represents either the unjustified transfer of wealth from those dependent on the acquired firm. . . to its

For concrete recent evidence that specifically *U.S.* firms in the early 2020's do not maximize profit, consider the manufacturing of digital semiconductor computer chips.[54] U.S. companies such as Texas Instruments, IBM, Motorola, and Intel were at one time leading manufacturers of these chips. Now, all of these but Intel have dropped out,[55] and Intel's technology is so far behind that it is contemplating[56] obtaining chips from its main chip fabrication rival, Taiwan Semiconductor Manufacturing Company (TSMC). TSMC was founded recently, in 1987,[57] by Morris Chang, who in 1983 was passed over for promotion to the CEO of Texas Instruments.[58] How has TSMC come to dominate (together with Samsung) the production of these cutting-edge chips, while all U.S. firms but one have given up, and the remaining one may partially give up? Together with a co-author, Clayton Christensen, formerly the Kim B. Clark Professor of Business Administration at the Harvard Business School and the recipient of eight honorary doctoral degrees,[59] explained:[60]

> Intel is the only significant U.S. semiconductor company that still makes its own chips. If you measure profitability using *return on assets*, the other companies are much more profitable, for a simple reason: Outsourcing fabrication to contractors like Taiwan Semiconductor Manufacturing Company (TSMC) reduces the denominator in that ratio. In 2009 Clay[ton] Christensen interviewed Morris Chang, founder of TSMC, about this phenomenon. Chang had been second-in-command at one of the most powerful semiconductor companies in America, Texas Instruments, before he returned to his native Taiwan and founded

shareholders, or valuation errors by the stock market.')."

[54]We are not discussing the much cheaper analog semiconductor chips; for the distinction, see Matt Franz, "The Semiconductor Industry from 10,000 Feet," available at https://www.eaglepointcap.com/blog/the-semiconductor-industry-from-10000-feet.

[55]See Debby Wu, Sohee Kim and Ian King, *Why the World Is Short of Computer Chips, and Why It Matters*, THE WASHINGTON POST (2021), available at https://www.washingtonpost.com/business/why-the-world-is-short-of-computer-chips-and-why-it-matters/2021/09/23/4943357a-1c2d-11ec-bea8-308ea134594f_story.html.

[56]See Wikipedia, *Semiconductor Industry*, available at https://en.wikipedia.org/wiki/Semiconductor_industry.

[57]See Wikipedia, *TSMC*, available at https://en.wikipedia.org/wiki/TSMC.

[58]*Supra* note 54.

[59]See Wikipedia, *Clayton Christensen*, available at https://en.wikipedia.org/wiki/Clayton_Christensen.

[60]Clayton M. Christensen and Derek van Bever, *The Capitalist's Dilemma*, 92/6 (June) HARVARD BUSINESS REVIEW 60 (2014) at 68. Emphasis added.

> TSMC. At the time of this interview, TSMC was making more than half of all semiconductor circuits in the world.
>
> Clay said to Chang, "Every time a new customer outsources to you, he peels assets off of his balance sheet, and in one way or another puts those assets on your balance sheet. You both can't be making the right decision." "Yes, if you measure different things, both can be right," Chang replied. "The Americans like ratios, like RONA [Return on Net Assets], EVA [Economic Value Added], ROCE [Return on Capital Employed], and so on. Driving assets off the balance sheets drives the ratios up. I keep looking. But so far I have not found a single bank that accepts deposits denominated in ratios. Banks only take currency.
>
> "There is capital everywhere," Chang continued. "And it is cheap. So why are the Americans so afraid of using capital?"

Christensen and van Bever explain the "orthodoxy of new finance":[61]

> Because they were taught to believe that the efficiency of capital was a virtue, financiers began measuring profitability not as dollars, yen, or yuan, but as ratios like RONA (return on net assets), ROIC (return on invested capital), and IRR (internal rate of return). These ratios are simply fractions, comprising a numerator and a denominator, but they gave investors and managers twice the number of levers to pull to improve their measured performance. To drive RONA or ROIC up, they could generate more profit to add to the numerator, of course. But if that seemed daunting, they could focus on reducing the denominator—outsourcing more, wiping more assets off the balance sheet.

Christensen and van Bever paint with a broad brush, and I would not want to endorse their sentiment that[62] "We have regressed from the decades when Drucker and Levitt urged us. . . to remember that the point of a business is to create a customer." The interest rate is a financial ratio, and it is critically important to anyone, external or internal to the firm, who has to allocate credit. But maximizing "return on (net) assets," in particular, is not equivalent to maximizing profit, and it leads to an obsession with wiping assets off of the balance sheet in order to make that financial ratio look good. An example is Vizio, which as of 2020 was the second-largest seller of flat-screen TVs in the U.S. and had revenue of $2 billion, and which has only

[61] *Supra* note 60 at 64.

[62] *Supra* note 60 at 68.

527 employees—fewer than the enrollment of some elementary schools.[63] The proper pecuniary goal of a company, accepting the traditional neoclassical point of view, is to maximize profit,[64] not to maximize the ratio of profit to net assets, but apparently large parts of the U.S. corporate sector are currently being controlled by people who do not know this. Christensen and van Bever point out a completely unsurprising consequence:[65] "For nearly a decade, the actual returns of all VC[venture capital]-backed investments, which were promised to be at least 25%, have totaled up to zero every year." Meanwhile, foreign companies like TSMC, which are not afraid to own capital equipment and have their own work force because they care not about financial ratios but about profit, flourish—as shown by TSMC's breaking ground on a $12 billion chip fabrication plant in Arizona and Samsung's plan to build a $17 billion chip fabrication plant outside of Austin, Texas.[66] There are certainly other authors who agree with Christensen and

[63] See Wikipedia, *Vizio*, available at https://en.wikipedia.org/wiki/Vizio, and Jerry Davis, *Capital Markets and Job Creation in the 21st Century*, CENTER FOR EFFECTIVE PUBLIC MANAGEMENT AT BROOKINGS (2015), available at https://www.brookings.edu/wp-content/uploads/2016/07/capital_markets.pdf.

[64] As evidence of this, there is not a single occurrence of phrases "rate of return" or "return on" in the best-known Ph.D.-level microeconomic theory textbook, a 981-page tome whose authors all hold or once held endowed professorships at Harvard or at MIT: ANDREU MAS-COLELL, MICHAEL D. WHINSTON, AND JERRY R. GREEN, MICROECONOMIC THEORY (1995).

[65] *Supra* note 60 at 66.

[66] *TSMC says has begun construction at its Arizona chip factory site*, reuters.com, June 1, 2021, https://www.reuters.com/technology/tsmc-says-construction-has-started-arizona-chip-factory-2021-06-01/; *Samsung says it will build $17B chip factory in Texas*, npr.org, November 24, 2021, https://www.npr.org/2021/11/24/1058770506/samsung-says-it-will-build-17b-chip-factory-in-texas.

For further explanation, see Leo Sadovy, *When Finance Informs the Factory*, INDUSTRY WEEK (2015), available at https://www.industryweek.com/finance/article/22008169/when-finance-informs-the-factory. Excerpts:

> Clayton Christensen, professor at the Harvard Business School and author of *The Innovator's Dilemma,* places the blame for the situation highlighted by the article on what Christensen calls the "Church of Finance." [. . .]
>
> With the Church of Finance and their capital efficiency metrics and ratios, it is no longer sufficient to make a ton of cash and call it a day. Our financial market-driven economy is obsessed with efficiency ratios—return on assets, return on capital employed, and return on equity.
>
> Unlike a scalar metric such as cash, ratios can be affected by changes in either their numerator or their denominator. And, as most experienced business and finance people know, it's a lot easier to shrink the asset-based denominator (which would include inventory) than to increase the revenue-dependent numerator.

van Bever's diagnosis.[67]

Christensen and van Bever's propose an explanation for *why* the goal of U.S. business leaders has been perverted from profit to financial ratios. They write:[68]

> In our view the crux of the problem is that investments in different types of innovation affect economies (and companies) in very different ways—but are evaluated using the same (flawed) metrics. Specifically, financial markets—and companies themselves—use assessment metrics that make innovations that eliminate jobs more attractive than those that create jobs. We'll argue that the reliance on those metrics is based on the outdated assumption

---

> Our desire to satisfy Wall Street has resulted in a lopsided, out-of-balance approach to investments in innovation. Over the past several decades, the emphasis has been on reducing that denominator, decreasing the use of cash, capital, assets and equity, to the detriment of investment in disruptive innovation that could affect the numerator.
>
> Christensen calls this the "Capitalist's Dilemma". . . . This unhealthy focus on the ratios, and especially on the denominator, on capital efficiency, has led to consequences entirely predictable from the model: fewer jobs, excess capital and uninvested cash. . . .
>
> When the Church of Finance runs the economy, the result is the jobless recovery the U.S. is currently experiencing.

[67] See William Levinson, *When Finance Runs the Factory*, INDUSTRY WEEK (2014), available at `https://www.industryweek.com/leadership/article/21963919/when-finance-runs-the-factory`. He writes: "The rest of this article will focus on why financial metrics that are mandatory for financial reports and tax returns have no place in operational decision making," and has sections entitled "When Labor and Overhead Are Not Costs," "Book Value Is Meaningless Except to the IRS and SEC" ("Depreciation, return on investment (ROI) and return on assets (ROA) are incentives to not replace aging machinery. . . "), and "When Cheap Offshore Labor is Expensive."

In 2014, Michael Sekora (Michael Sekora, *Reviving U.S. Manufacturing is the Wrong Goal to Set to Improve The Economy*, FORBES (2014), available at `https://www.forbes.com/sites/beltway/2014/02/04/why-reviving-u-s-manufacturing-is-the-wrong-goal-to-set-to-improve-the-economy/?sh=4d4305a1763e`) wrote, "The disease killing America's economic health is financial-based planning, and one of the symptoms of this ongoing disease is the loss of the U.S. manufacturing base. The total reliance of American companies, governments and academic institutions on financial-based planning is what caused them to divest the country of its manufacturing base and blinded them to the fact that this would cause a major shift in economic might from the U.S. to China and others."

For the dire consequences of using bad financial metrics, see the following article by the Raphael Dorman-Helen Starbuck Professor of Political Science at MIT: Suzanne Berger , *How Finance Gutted Manufacturing*, BOSTON REVIEW (2014), available at `https://bostonreview.net/forum/suzanne-berger-how-finance-gutted-manufacturing`.

[68] *Supra* note 60 at 62.

> that capital is, in George Gilder's language ["Every economic era is based on a key abundance and a key scarcity"], a "scarce resource" that should be conserved at all costs. But, as we will explain further, capital is no longer in short supply—witness the $1.6 trillion in cash on corporate balance sheets—and, if companies want to maximize returns on it, they must stop behaving as if it were. We would contend that the ability to attract talent, and the processes and resolve to deploy it against growth opportunities, are far harder to come by than cash. The tools businesses use to judge investments and their understanding of what is scarce and costly need to catch up with that new reality.

However, one should be cautious in asserting that capital—or, more precisely, credit—was ever in short supply. From a borrower's point of view, it is true that credit can be in short supply, because the interest rate may be too high or banks may be unwilling to lend at the current interest rate (this is called "credit rationing"). But credit is never in short supply from the bank's point of view, in terms of being difficult to create, because banks and other lending institutions know that credit can be costlessly created at any moment by making offsetting entries on both sides of a balance sheet, which is what banks do when they make a loan.[69] (Credit certainly may be imprudent to create, and this can depend on whether the bank's capitalization is in short supply). In this sense, commercial banks are not as different as one might think from the Federal Reserve, which on the spur of the moment came up with roughly $80 billion to bail out AIG in 2008.[70] The incorrect notion that credit is a scarce resource may come from an ideological attempt to defend the income of lenders and bankers: workers deserve payment because labor is scarce and effortful, suppliers of other physical inputs deserve payment because those inputs are scarce and effortful to create, so unless credit was thought difficult to create, questions might be raised about why its suppliers deserved to be paid. (Its suppliers deserved to be paid to the extent that they

[69]This is not what most economics textbooks teach—they teach that banks are intermediaries, and that loans come from bank deposits—but it is the clearly stated (and correct) view of, among others, the Bank of England. See Michael McLeay, Amar Radia and Ryland Thomas, *Money Creation in the Modern Economy*, March QUARTERLY BULLETIN OF THE BANK OF ENGLAND 1 (2014). Available at `https://www.bankofengland.co.uk/quarterly-bulletin/2014/q1/money-creation-in-the-modern-economy`.

[70]See William Greider, *Dismantling the Temple: How to Fix the Federal Reserve*, 289/4 THE NATION 11 (2009). For House Speaker Nancy Pelosi's comment on the ability of the Federal Reserve to do that, see minutes 9:25 to 10:00 of her June 18, 2009 news conference at `https://www.c-span.org/video/?287114-1/house-speaker-news-conference`.

extend credit to worthy endeavors and deny credit to unworthy endeavors. This makes the dismal recent returns to U.S. venture capital telling: it is precisely what one would expect if VC has been using bad metrics to allocate credit.)

A better explanation for why financial ratios have such misplaced importance starts with the observation that for someone who allocates credit, such as a banker, it is necessary to use a ratio to determine the return to extending credit, because both the outflow and inflow are fundamentally in dollars (while assets, for example, are not fundamentally in dollars, they are fundamentally in other units such as tons of steel-making capacity or thousands of square feet of office space, and they have to be translated into dollars). In the nineteenth century, credit was mostly used to purchase capital equipment ("assets"), so in that era

$$\frac{\text{return}}{\text{credit}} \approx \frac{\text{return}}{\text{assets}};$$

and since the left-hand side is an appropriate criterion to use when allocating or extending credit, the right-hand side would be acceptable. However, there can no longer be a presumption that firms use credit mostly to buy assets—certainly that is not true of software firms, nor of other firms where there is considerable intellectual capital. For such firms, "return on employee salaries" makes no less sense than "return on assets." This makes "return on assets" an uninformative criterion today, even for lenders. For anyone whose job does not involve allocating credit, "return on assets" was never an appropriate metric to judge firm performance.

It follows that antitrust economists should empirically measure the extent to which a firm's pecuniary motivation is profit, and antitrust decisions should not expand the resource-allocation reach of managers whose pecuniary goal is found to be maximization of return on assets instead of maximization of profit. To do otherwise could threaten the long-run sustainability of the firm by destroying the critical human capital that constitutes the firm's knowledge base, all for no true gain in profit.

C. Firms' pecuniary motivations: multi-period settings. When considering multi-period settings, incorporating both short-run and long-run considerations, economists theorize that firms maximize the "net present value of profit" ("NPV"), meaning that if profits in periods 1, 2, 3, ... are $\pi_1, \pi_2, \pi_3, \ldots$, and if r is a discount rate, then firms maximize

$$\text{NPV} = \frac{\pi_1}{(1+r)^1} + \frac{\pi_2}{(1+r)^2} + \frac{\pi_3}{(1+r)^3} + \cdots . \tag{7}$$

NPV is measured in dollars; it is not a ratio. Firms practice socially-destructive "short-termism" if they use a discount rate which is higher than would be socially appropriate. For example, suppose management has a choice between: (a) operating the firm in a long-run, sustainable fashion, making an initial investment of $10 million and obtaining returns of $2 million for each of the next 20 years; or (b) operating the firm just for short-term payback, making an initial investment of $10 million and obtaining a return of $15 million for the next year and then going out of business. Using (7), it turns out that for a typical discount rate such as 5% or 3% or even 8%, management plan (a) gives a higher NPV than management plan (b). However, if management chooses to use a discount rate above approximately 14.1%, the "predatory" management plan (b) gives a higher NPV. Economists in antitrust cases should find out what discount rate the involved firms are accustomed to using, and courts should look askance at firms run by managers using inappropriately high discount rates.

The rate of return which managers advertise to potential investors can be a signal of the discount rate those managers use in making management decisions. Above, we saw that Christensen and van Bever report that venture capitalists think they will earn at least a 25% return.[71] This makes it likely that venture capitalists routinely use a discount rate of at least 25% when making management decisions. This is much higher than would be socially optimal, since realistic market returns are much less than 25% on average.

Managers' inappropriate obsession with financial ratios extends to multi-period decision-making, where they often use flawed metrics such as "internal rate of return" or "accounting rate of return," instead of NPV, to make management decisions.[72] Economists in antitrust cases should find out whether any management team is making multi-period decisions using

[71] *Supra* note 65.

[72] For the superiority of the net present value criterion over criteria such as Internal Rate of Return, see STEPHEN A. ROSS, RANDOLPH W. WESTERFIELD, AND JEFFREY F. JAFFE, CORPORATE FINANCE (THIRD EDITION) (1993) at 155 ("... we believe that the NPV approach is the best one for evaluating capital budgeting projects. ..."), Brian Balyeat and Julie Cagle, *MIRR: The Means to an End? Reinforcing Optimal Investment Decisions Using the NPV Rule*, 41/1 JOURNAL OF FINANCIAL EDUCATION 90 (2015) at 99 ("... the finance discipline is fairly *univocal* on the NPV rule being the decision technique most consistent with shareholder wealth maximization"), and Carlton L. Dudley, Jr., *A Note on Reinvestment Assumptions in Choosing between Net Present Value and Internal Rate of Return*, 27/4 THE JOURNAL OF FINANCE 907 (1972) at 913, where the equivalence of maximizing NPV and maximizing *terminal* value (or "future value"), which is fundamentally what is important even though terminal value is usually ignored, is discussed. For extensions of the NPV rule to incorporate optionality, see Stephen A. Ross, *Uses, Abuses, and Alternatives to the Net-Present-Value Rule*, 24/3 FINANCIAL MANAGEMENT 96 (1995).

a ratio (a percentage) instead of using NPV, because such decisions are mistakes which help no one.[73]

8. Beyond the Firm's Owners, Managers, and Customers: Towards a New Brandeisian Econometrics of Antitrust

Shapiro[74] writes that the New Brandeis scholars "are highly skeptical of the role of economics and expertise in antitrust." As shown above, they have

[73]By now, economists are agreed that NPV, not internal rate of return ("IRR") nor the "accounting rate of return," is the appropriate criterion for multi-period decision making, but because business management teams prefer to think in terms of ratios, there is a huge literature attempting to either defend the use of IRR or the "accounting rate of return"—which is futile—or trying to find some new ratio which would be equivalent to NPV. I base these comments not only on the sources given in *supra* note 72 but also on the masterful, though highly mathematical, treatment of this topic in Carlo Alberto Magni, *Capital Depreciation and the Underdetermination of Rate of Return: A Unifying Perspective*, 67 JOURNAL OF MATHEMATICAL ECONOMICS 54 (2016). Economists might be familiar with the early 1980's debate between Franklin Fisher and his critics on the appropriateness of "accounting rates of return," prompted by Fisher's testimony in *U.S. v. IBM* (69 Civ. 200, U.S. District Court, Southern District of New York) (see Franklin M. Fisher and John J. McGowan, *On the Misuse of Accounting Rates of Return to Infer Monopoly Profits*, 73/1 THE AMERICAN ECONOMIC REVIEW 82 (1983) and Franklin M. Fisher, *The Misuse of Accounting Rates of Return: Reply*, 74/3 THE AMERICAN ECONOMIC REVIEW 509 (1984) and the preceding articles involved in the debate, which are cited in his bibliography). Fisher and McGowan were right to criticize accounting rates of return, but their advocacy of IRR (*id.* at 509) was mistaken. In a footnote (*id.* note 4) Fisher claimed that "Contrary to what some of my correspondents appear to believe," he was "not recommending internal rate of return calculations as a substitute for present value maximization," but on the same page he wrote that "The economic rate of return [i.e., IRR] is the magnitude which gives the signal for entry or exit of resources," which certainly confuses matters, and much of Fisher and McGowan's paper is based on their definition (A1), which is an internal rate of return. Net present value is actually older than IRR (Magni *op. cit.* at 55 attributes IRR to a 1935 paper of Kenneth Boulding), and NPV was correctly used to describe capital valuation almost a century ago in Harold Hotelling, *A General Mathematical Theory of Depreciation*, 20/151 JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION 340 (1925); for a more streamlined treatment see Gabriel A. Lozada, *Resource Depletion, National Income Accounting, and the Value of Optimal Dynamic Programs*, 17 RESOURCE AND ENERGY ECONOMICS 137 (1995). Fisher and McGowan's criticism of "depreciation schedules anything like those used by real-life firms to optimize after-tax profits given IRS rules or those schedules used for nontax purposes" is well-taken, and the heart of the problem with "accounting rates of return" is their *ad hoc* depreciation schedules. On the other hand, Magni invents a ratio called the "average IRR" that is based on accounting data and that is equivalent to NPV (see his Proposition 14, at 61); there can be *no* objection to using it to make decisions. Note that "return on assets" attacked in the previous subsection of this paper is not the same thing as "accounting rates of return."

[74]*Supra* note 1.

reason to be so. Yet our critique up to this point has accepted the exclusive focus which empirical Industrial Organization economics has put on consumer demand and on conditions inside the firm, the former a reflection of the "consumer welfare" standard and the latter a reflection of its "producer surplus" counterpart, a framing which the New Brandeis School objects to as being overly narrow. There is an element of disingenuousness when Shapiro writes that the consumer welfare standard[75] "has outlived its usefulness and should be retired [. . . lest. . .] the public [be. . .] all too easily be left with the impression that antitrust is just about final consumers and ignores the interests of workers, farmers, or business customers." After all, the reason the public has been left with the impression that antitrust is just about final consumers and ignores the interests of workers, farmers, or business customers is because until now antitrust econometrics *has* been just about final consumers and *has* completely ignored the interests of workers, farmers, or business customers. (Antitrust econometrics has not even been about everything final consumers care about, including durability or warranty provisions or customer service; antitrust econometrics has instead solely focused on price and quantity.) In antitrust, paired with the ubiquitous econometric analysis of consumer demand, going forward there should be a similarly ubiquitous econometric analysis of the impact of mergers and acquisitions on workers and on the labor market.

For that matter, going forward there should be econometric analyses of the impact of mergers and acquisitions on farmers, when the topic is agriculture, and there should be econometric analysis estimating the social cost of any increase in water and air pollution, if increased industry concentration affects the local or global environment; and there should be separate econometric analyses of the merger on each one of the affected local economies, not just an aggregate study on the overall national economy.

Shapiro[76] tells us that "antitrust developed the mantra that 'antitrust protects competition, not competitors'," but one would search in vain for a passage in a modern textbook on welfare economics instructing the reader that consumers have a privileged place compared to others. Indeed, welfare economists often adopt the "anonymity" assumption, which holds that all individuals are treated equally.[77] If an economic development helps consumers and hurts competitors, we need an econometric study to determine

[75] *Id.* at 38–39.

[76] *Id.* at 38.

[77] Christian List, *Social Choice Theory* (2013). In Edward N. Zalta, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY at 2.2. Available at `https://plato.stanford.edu/archives/win2013/entries/social-choice`.

whether that development is or is not in the social interest, not an unscientific mantra—are any mantras scientific?—which tells us *a priori* that it is in the social interest.[78] The competitors who would be hurt by the development could likely provide the econometrician with suggestions for quite a few socioeconomic impacts to measure, including the effect of going out of business on their and their workers' future earnings, suicide and divorce rates (and the economic consequences of increasing suicides and divorces), the rate at which their children's human capital is going to be developed, and the response of crime and opioid addiction statistics, with their resulting economic impacts.

By making dissemination of misinformation and genocidal propaganda easier, Facebook's economies of scale have generated externalities in the form of, for example, ethnic cleansing of the Rohingya people in Myanmar/Burma.[79] Can we have econometric estimates of the social cost of those activities? And by the way, lest one think Section 6's insistence on the difference between equivalent variation and compensating variation is a trivial, technical economic matter, the willingness and ability of a Rohingya peasant to pay for keeping a family member from being killed is going to be very much less than that peasant's willingness to accept compensation for that family member's death, so which of the two valuation methods is the econometrician going to try to measure? And what is the econometric assessment method to value a young child's loss of his or her Rohingya parent, given the different earnings future and life course which the child would likely experience with or without the parent and with or without the genocidal environment?

Where are the econometric models quantifying the effect of increased market concentration on campaign contributions, lobbying expenditures, media disinformation campaigns, and partisan think tank support, quantifying the then-consequent changes in laws and regulations, and quantifying the possible long-run resulting increase in income inequality? Shapiro[80] writes,

---

[78]The notion that it *must* be in the social interest is probably based on the idea that the development is a Potential Pareto Improvement, but that itself is a controversial criterion: see Mark Glick and Gabriel A. Lozada, *The Erroneous Foundations of Law and Economics*, INSTITUTE FOR NEW ECONOMIC THINKING, WORKING PAPER 149 (2021), available at `https://www.ineteconomics.org/research/research-papers/the-erroneous-foundations-of-law-and-economics`.

[79]See BBC News, "Facebook admits it was used to 'incite offline violence' in Myanmar," Nov. 6, 2018, available at `https://www.bbc.com/news/world-asia-46105934`

[80]*Supra* note 1 at 42.

> The Populists [i.e., New Brandeisians] implicate lax antitrust as the central cause of many of our social and economic problems, while I see other public policy failures—including weak voter-protection and anti-corruption laws, inadequate protections for workers, highly unequal access to education and health care, and a tax system that contains many regressive elements—as the central culprits.

Shapiro's error is in thinking that market power has no impact on these "other public policy failures." On the contrary, increased market power enables those who possess it to achieve increased success in having government enact their public policy desires. Accordingly, the New Brandeis School realizes that we need a much broader econometrics of the harm done by market concentration.

In addition to a much broader econometrics of the harm done by market concentration, we also need an econometric understanding of broader possible *remedies* for such concentration. Shapiro[81] writes:

> Breaking up monopolies and oligopolies on a "no-fault" basis would be fighting against powerful, underlying economic forces associated with economies of scale and scope that are a pervasive feature of advanced economies around the world.

When an industry is characterized by strong "economies of scale and scope," two possibilities beyond Shapiro's dichotomy—breaking up its firms or leaving them alone—are to put the firms' activities under tight, never-ending government regulation, as is now done for public utilities such as electricity and natural gas, or simply nationalizing the firms. Both these options retain the economics of scale and scope. Econometric study of nationalized businesses is uncommon, but Szarzec et al. show more good than bad in state-owned enterprises in countries with good institutions.[82] Another policy suggestion would be to mandate worker representation on companies' boards of directors, as is not uncommon in Europe, Germany being the best-known example.[83] Again, this policy is understudied econometrically, but

[81] *Id.* at 43.

[82] Katarzyna Szarzec, Ákos Dombi, and Piotr Matuszaka, *State-owned Enterprises and Economic Growth: Evidence from the Post-Lehman Period*, 99/105490 ECONOMIC MODELLING 1 (2021).

[83] See Wikipedia, *Worker Representation on Corporate Boards of Directors*, available at `https://en.wikipedia.org/wiki/Worker_representation_on_corporate_boards_of_directors`. For remarks on a related proposal by Sen. Elizabeth Warren, see Raffaella Sadun, *Worker Representation on Boards Won't Work Without Trust*, HARVARD BUSINESS

Kim et al. find that it helps some workers and hurts neither shareholders nor anyone else.[84]

In short, there could be such a thing as a New Brandeisian econometrics. But whether that comes about or not, in principle, one should not limit consideration solely to phenomena that can be measured. Keep in mind the "McNamara fallacy"[85]:

> The first step is to measure what can be easily measured. This is okay as far as it goes. The second step is to disregard that which cannot be measured, or give it an arbitrary quantitative value. This is artificial and misleading. The third step is to presume that what cannot be measured really isn't important. This is blindness. The fourth step is to say that what can't be easily measured really doesn't exist. This is suicide.

Conclusion

Particularly when assuming a representative consumer, Industrial Organization econometricians make many assumptions whose validity is questionable. For this reason, econometricians in other fields of economics are increasingly moving away from representative-consumer modeling. However, doing so requires comprehensive, household-level panel data, which antitrust investigations usually lack. This leaves antitrust econometrics on a questionable footing, often obscured by its mathematical sophistication.

However, even if an econometric analysis is done in a compelling way, it obscures more than it illuminates if it omits major effects of market concentration on any affected people. Workers, competitors, the competitors' workers, pollution victims, neighbors in affected communities—all are completely ignored in the standard Industrial Organization approach to antitrust analysis. There is no justification, in economics or outside of economics, for such glaring omissions. Similarly, the proposition that firms maximize profit (or the net present value of profit, with a discount rate that makes the

---

REVIEW (2018), available at https://hbr.org/2018/08/worker-representation-on-boards-wont-work-without-trust.

[84]E. Han Kim, Ernst Maug, and Christoph Schneider, *Labor Representation in Governance as an Insurance Mechanism*, 22/4 REVIEW OF FINANCE 1251 (2018). See also Alex Edmans, *Do Workers on Boards Protect Jobs in a Downturn?* (no date). Available at https://alexedmans.com/blog/responsible-business/the-causal-effect-of-workers-on-boards.

[85]Coined by Daniel Yankelovich, who named it after U.S. President Lyndon Johnson's Secretary of Defense Robert McNamara. Source: "ADAM SMITH" [PSEUDONYM OF GEORGE J. W. GOODMAN], SUPERMONEY (1972) at 290.

future meaningful rather than insignificant) should be subject to econometric scrutiny, not simply accepted without proof, because if managers are sacrificing profit for social goods, or on the other hand are maximizing irrelevant indicators or only short-term metrics, these behaviors should have a bearing on antitrust decisions.

Appendix

Proof of Proposition 1. Following Muellbauer,[86] let $w_i = (p_i d_i)/I$ be the value share of good i, where we only consider one household and we suppress the h superscript. From Muellbauer's Theorem 3, using slightly different notation, there exist functions $A'_i(\mathbf{p})$, $B'_i(\mathbf{p})$, and $v(I, \mathbf{p})$ such that[87]

$$w_i(I, \mathbf{p}) = v(I, \mathbf{p}) A'_i(\mathbf{p}) + B'_i(\mathbf{p}) \tag{8}$$

$$w_j(I, \mathbf{p}) = v(I, \mathbf{p}) A'_j(\mathbf{p}) + B'_j(\mathbf{p}) \,. \tag{9}$$

Solving (9) for $v(I, \mathbf{p})$ and substituting it into (8) leads to

$$\begin{aligned} w_i(I, \mathbf{p}) &= \frac{w_j(I, \mathbf{p}) - B'_j(\mathbf{p})}{A'_j(\mathbf{p})} A'_i(\mathbf{p}) + B'_i(\mathbf{p}) \\ &= \frac{A'_i(\mathbf{p})}{A'_j(\mathbf{p})} w_j(I, \mathbf{p}) + B'_i(\mathbf{p}) - \frac{B'_j(\mathbf{p})\, A'_i(\mathbf{p})}{A'_j(\mathbf{p})} \end{aligned}$$

which for simplicity we will write as

$$w_i(I, \mathbf{p}) = A''_{ij}(\mathbf{p})\, w_j(I, \mathbf{p}) + B''_{ij}(\mathbf{p}) \,.$$

Substituting in the definition of value shares leads to

$$\begin{aligned} \frac{p_i\, d_i(I, \mathbf{p})}{I} &= A''_{ij}(\mathbf{p})\, \frac{p_j\, d_j(I, \mathbf{p})}{I} + B''_{ij}(\mathbf{p}) \,, \text{ so} \\ d_i(I, \mathbf{p}) &= A''_{ij}(\mathbf{p})\, \frac{p_j}{p_i}\, d_j(I, \mathbf{p}) + \frac{B''_{ij}(\mathbf{p})}{p_i}\, I \,, \end{aligned}$$

which for simplicity we can write as

$$d_i(I, \mathbf{p}) = A_{ij}(\mathbf{p})\, d_j(I, \mathbf{p}) + B_{ij}(\mathbf{p})\, I \,.$$

Take the partial derivative of both sides with respect to I. ∎

[86] *Supra* note 8.

[87] Here v is not an indirect utility function; Muellbauer (at 530) interprets it as being the value share for an arbitrary (say, the first) commodity. Muellbauer uses an upper-case V to denote indirect utility functions, as in his equation (12).

Proof of Proposition 2. Differentiating both sides of (3) with respect to I^h,

$$\frac{\partial^2 d_i^h(I^h, \mathbf{p})}{\partial (I^h)^2} = A_{hij} \frac{\partial^2 d_j^h(I^h, \mathbf{p})}{\partial (I^h)^2} \quad \text{for each } h \text{ and for all goods } i \neq j. \tag{10}$$

This proves the proposition's second sentence.

A function is linear if and only if its second derivative is zero.[88] Therefore, wherever d_j^h is linear, the right-hand side of (10) is zero, so the left-hand side of (10) must be zero as well, making d_i^h is linear. ∎

Proof of Corollary. A constant function is linear; use Proposition 2. ∎

Symmetry, re-expressed. Deaton and Muellbauer[89] state our equation (4) as their equation ([2.]4.6), but at 76, their equation ([3.]4.11) expresses symmetry in a completely different way, without any explanation, and the latter expression is the one used by most later authors. Ulrick[90] has an explanation; here is a different one, which is more elementary in some respects. We use Deaton and Muellbauer's notation for income and amount consumed instead of ours.

Proposition 3. *If w_i is the value share of consumption of the ith commodity, p_i is the price of the ith commodity, and x is income, the AIDS demand system is*

$$w_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log\{x/P\}$$

where by definition P satisfies

$$\log P = \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2} \sum_j \sum_k \gamma_{kj} \log p_k \log p_j \,.$$

In the AIDS demand system, the condition that the Slutsky Substitution Matrix be symmetric is equivalent to the condition that $\gamma_{ij} = \gamma_{ji}$.

Proof. This description of the AIDS demand system comes from Deaton and Muellbauer's[91] equations (8) and (9). Their equation (6) can be written as

$$\frac{p_i q_i}{x} = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i u \beta_0 \prod_k p_k^{\beta_k}$$

where q_i is quantity demanded and x is income (or expenditures). Therefore

$$q_i = \frac{x}{p_i}\alpha_i + \frac{x}{p_i} \sum_j \gamma_{ij} \log p_j + \frac{x}{p_i} \beta_i u \beta_0 \prod_k p_k^{\beta_k} \,.$$

[88] A mathematician would use "affine" wherever we use "linear."

[89] *Supra* note 14 at 45.

[90] *Supra* note 6 at 132.

[91] *Supra* note 20.

For example, if $n = 3$ then

$$q_i = \frac{x}{p_i}\alpha_i + \frac{x}{p_i}(\gamma_{i1}\log p_1 + \gamma_{i2}\log p_2 + \gamma_{i3}\log p_3) + \frac{x}{p_i}\beta_i u\beta_0 p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3}$$

and

$$q_j = \frac{x}{p_j}\alpha_j + \frac{x}{p_j}(\gamma_{j1}\log p_1 + \gamma_{j2}\log p_2 + \gamma_{j3}\log p_3) + \frac{x}{p_j}\beta_j u\beta_0 p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3}\,.$$

Note that since

$$\frac{\partial}{\partial p_1} p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3} = \beta_1 p_1^{\beta_1 - 1} p_2^{\beta_2} p_3^{\beta_3} = \frac{\beta_1}{p_1} p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3}$$

we can write by analogy

$$\frac{\partial}{\partial p_i} p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3} = \frac{\beta_i}{p_i} p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3}\,,$$

which makes it easy to calculate the following derivatives for $i \neq j$:

$$\begin{aligned}
\frac{\partial q_i}{\partial p_j} &= \frac{x}{p_i}\gamma_{ij}\frac{1}{p_j} + \frac{x}{p_i}\beta_i u\beta_0\frac{\beta_j}{p_j} p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3} \\
&= \frac{x\gamma_{ij}}{p_i p_j} + \frac{x\beta_i\beta_j\beta_0 u}{p_i p_j} p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3} \qquad (11) \\
\frac{\partial q_j}{\partial p_i} &= \frac{x}{p_j}\gamma_{ji}\frac{1}{p_i} + \frac{x}{p_j}\beta_j u\beta_0\frac{\beta_i}{p_i} p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3} \\
&= \frac{x\gamma_{ji}}{p_i p_j} + \frac{x\beta_i\beta_j\beta_0 u}{p_i p_j} p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3}\,. \qquad (12)
\end{aligned}$$

The symmetry condition on the Slutsky Substitution Matrix is that the left-hand side of (11) be equal to the left-hand side of (12) (since these q's are Hicksian (compensated) demand curves, not Marshallian demand curves). Clearly this holds if and only if $\gamma_{ij} = \gamma_{ji}$.

To extend this to Marshallian demand curves, in Deaton and Muellbauer's equation (4), replace the expenditure function $c(u, p)$ with income (or expenditure) x and rearrange to obtain, for $n = 3$,

$$\beta_0 u p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3} = \log x - \left(\alpha_0 + \sum_{k=1}^{3}\alpha_k \log p_k + \frac{1}{2}\sum_{k=1}^{3}\sum_{j=1}^{3}\gamma_{ij}^{*}\log p_k \log p_j\right). \qquad (13)$$

Replacing the $\beta_0 u p_1^{\beta_1} p_2^{\beta_2} p_3^{\beta_3}$ portions of (11) and (12) with the right-hand side of (13) drops u and introduces x, thus turning the expressions into their Marshallian form.[92]

The extension to other values of n is straightforward. ∎

[92] Usually, Marshallian demand curves do not have the property that $\partial q_i/\partial p_j = \partial q_j/\partial p_i$, but this proof suggests that in the AIDS model, the Marshallian demand curves do have that property.

As Ulrick[93] also notes, in the first sentence of Proposition 3, the γ_{ij} constants are defined by Deaton and Muellbauer[94] as $\gamma_{ij} = \frac{1}{2}(\gamma^*_{ij} + \gamma^*_{ji})$, "hence, $\gamma_{ij} = \gamma_{ji}$ and symmetry holds. (When estimating the AIDS model, it is possible to obtain estimated parameters such that $\gamma_{ij} \neq \gamma_{ji} \ldots$.)"

[93] *Supra* note 6 at 132.

[94] Equation (7) of *supra* note 20.