

Paul Samuelson and Revealed Preference Theory*

D. Wade Hands
Department of Economics
University of Puget Sound
Tacoma, WA 98416
hands@pugetsound.edu
12,594 Words
Version 1.3

Final Pre-Publication Version
Published in: *History of Political Economy*, 46, 2014, 85-116.
DOI 10.1215/00182702-2398939

* I would like to thank Will Hansen and other members of the staff of the Duke University Rare Book, Manuscript, and Special Collections Library for the help I have received working with the Paul Samuelson Papers. I would also like to thank John Davis, Daniel Hausman, Kevin Hoover, and an anonymous referee for helpful comments on earlier versions of the paper. And finally, I would like to thank The Institute for New Economic Thinking (INET) for providing financial support through Grant IN01100003.

Abstract: Revealed preference theory is not a specific theory; it is a broad programmatic framework for analyzing choice behavior. Within this broad framework there are a number of different revealed preference theories (different versions of the program) – they all share common features, but there are also sharp differences. The diversity of revealed preference theory is not well understood, and one purpose of this paper is to improve our historical understanding of the field by examining this historical diversity. This history is valuable for its own sake, but also because it is relevant to recent debates about the methodological foundations of rational choice theory among experimental psychologists, behavioral economists, neuroeconomists, and others. The second purpose of this paper is to use material from the Paul Samuelson archives to help us understand how Samuelson, the originator of revealed preference theory, viewed his contribution to the program and how he evaluated the different versions of revealed preference theory. The paper will examine Das Paul Samuelson Problem: the question of whether Paul Samuelson changed his mind about the foundations (nature, significance, and purpose of) revealed preference theory over time.

Many writers have held the utility analysis to be an integral and important part of economic theory. Some have even sought to employ its applicability as a test criterion by which economics might be separated from the other social sciences. Nevertheless, I wonder how much economic theory would be changed if either of the two conditions above were found to be empirically untrue. I suspect very little.

(Samuelson, 1947 and 1983, p. 117)

Dangerous, I say, to insist on Central Messages. I have 3 different ones before breakfast most week days.

(Samuelson to Robert K. Merton, no date but early 1990s, Box 53)

0. Introduction

Paul Samuelson was twenty-three years old in 1938 when "A Note on the Pure Theory of Consumer's Behaviour" appeared in Economica. The paper, the origin of what later came to be called revealed preference theory, was one of Samuelson's very first publications and one of a group of his papers published that year on consumer choice theory (Samuelson 1938a, 1938b, 1938c, 1938d). Although Samuelson has consistently received credit for his ground-breaking work in revealed preference theory, the exact nature and significance of revealed preference has long been, and continues to be, a controversial question. For some, revealed preference is an alternative framework for the analysis of consumer choice, a replacement for the consumer choice (demand) theory – ordinal utility theory – that has dominated economics since the late 1940s. For others, revealed preference is a complement rather than a substitute for ordinal utility theory – simply a different, formally equivalent, way of expressing the standard theory. And as will be clear below, these two interpretations are not the only ones in the literature. For example, the interpretation I will call "contemporary revealed preference theory" sees revealed preference as not only equivalent to ordinal utility theory, but exhaustive of choice theory; Faruk Gul and Wolfgang Pesendorfer say it is "very similar to what can be found in a standard graduate textbook" (2008, p. 7), and Kenneth Binmore calls it the "official doctrine of neoclassical economics, enshrined in all respectable textbooks" (2009a, p.20).

One of the tasks of this paper will be to sort out these various versions of revealed preference theory. I will argue that revealed preference is not a single "theory," but rather is a broad research program – a family of theories – each with some shared features (family resemblances) but also a number of distinct differences. I will make the case that understanding the diversity within the revealed preference

program is necessary for understanding the various conflicting interpretations of what "it" is, as well as how both defenders and critics of revealed preference often end up arguing at cross purposes. Distinguishing these various interpretations requires an historical understanding of the revealed preference program: what contributions led to which later developments, and how the economists involved perceived the goals of, and issues surrounding, their particular subprogram. Samuelson himself played a substantial role in the development of many, but not all of, these various subprograms.

Samuelson's role leads to the second task of the paper – to try to situate Samuelson's interpretation of revealed preference within/among the various members of the revealed preference family. Did Samuelson support one, or a few, of the various approaches over others? Which later programs were most consistent with the original project of Samuelson (1938a)? And was Samuelson himself consistent? Did he change his mind about revealed preference, particularly between the time he was writing the original 1938 paper and his second round of contributions (Samuelson 1948a, 1950)? I will call this latter question the Paul Samuelson Problem (obviously with reference to the Adam Smith Problem: the extensive historical debate over whether Adam Smith changed his mind between The Theory of Moral Sentiments and The Wealth of Nations). The discussion of Samuelson's own view of revealed preference – what it was and whether it was stable – will involve both published works and materials from the Samuelson archives.

1. Varieties of Revealed Preference

Revealed preference is a broad research program in choice theory. All of the various versions of revealed preference share the idea that consistency conditions can be substituted for constrained optimization, and all versions originate, to some degree, in Samuelson's 1938 paper. The key revealed preference insight is that if the agent chooses A over B when both are available (A is revealed preferred to B) then consistent behavior requires that B would only be chosen if A is not available. There is a version of revealed preference that is quite abstract, not involving preference or individual agency at all, but rather choice functions defined over choice sets (Arrow 1959); there is (or was) a version of revealed preference that had nothing to do with agent choice at all, but rather revealed preference axioms, usually the "Weak Axiom of Revealed Preference," imposed on aggregate excess demand functions in Walrasian general equilibrium models¹; there are stochastic versions of revealed

¹ WARP on aggregate excess demand functions is sufficient for stability and uniqueness of the general equilibrium price vector. Wald (1951) was the first to use such a condition in a Walrasian model, but the key contribution was Arrow, Block, and Hurwicz (1959). See Arrow and Hahn (1971) or Weintraub (1991) for a detailed discussion of the literature. Just for the record, Samuelson never supported this version of

preference theory (e.g. Bandyopadhyay, Dasgupta, and Pattanaik 2004), and many others, but none of these are really the subject here. The subject here will be revealed preference approaches to consumer choice theory (demand theory): individual economic agents choosing bundles of commodities subject to a budget constraint. Consumer choice was the context for Samuelson's original work, and it has consistently been the main trunk of the revealed preference family tree. But restricting the discussion to consumer choice theory does not eliminate the diversity within the program; there are many different versions of revealed preference-based consumer choice theory, and this section will discuss the more important among them.

Samuelson 1938a is a clear and forcefully written paper. Its stated purpose is programmatic and unambiguous. The ordinal revolution of Hicks and Allen (1934), Pareto (1927), Slutsky (1934), and others had moved away from hedonistic and cardinal notions of utility, but according to Samuelson they had not gone far enough; the next step was the total elimination of the terms preference and utility from the theory of consumer choice. As Samuelson explained:

... despite the fact that the notion of utility has been repudiated or ignored by modern theory, it is clear that much of even the most modern analysis shows vestigial traces of the utility concept ... I propose, therefore, that we start anew in direct attack upon the problem, dropping off the last vestiges of the utility analysis. (1938a, pp. 61-62)

Early ordinal utility theorists had started from what they considered potentially observable phenomena – indifference curves (Pareto), price ratios at purchase points equal to the marginal rate of substitution (Hicks and Allen), and others – but since all of these foundational observables involved preference or utility in some way, Samuelson grounded his new theory elsewhere: in individual demand functions. As he explained:

I assume in the beginning as known, i.e. empirically determinable under ideal conditions, the amounts of n economic goods which will be purchased per unit time by an individual faced with the prices of these goods and with a given total expenditure. (ibid., p. 62)

revealed preference theory because it essentially reduced the Walrasian system to a model of one big consumer – a representative agent – which undermined the entire idea of competitive markets coordinating the actions heterogeneous agents (Samuelson 1955, pp. 499-500; Dorfman, Samuelson, and Solow, 1958, p. 368). A lesson perhaps for contemporary macroeconomics?

The initial givens were thus the individual demand functions $x_i = h_i(p_1, p_2, \dots, p_n, M)$ for all $i = 1, 2, \dots, n$, where p_i is the (competitive) price of good i , and M is money income. These functions were assumed to be single-valued and satisfy the budget constraint. In the original paper Samuelson also assumed they were differentiable, homogeneous of degree zero, and generated a non-vanishing Jacobian matrix, although in later work he demonstrated that not all of these restrictions were required for his main result.

The key to Samuelson's theory – what came to be called the "Weak Axiom of Revealed Preference" (WARP) – was a consistency condition on consumer demand functions (the prices and quantities purchased). The weak axiom said that if the consumer chose bundle x^0 at the prices p^0 when bundle x^1 was affordable – in the language that emerged later, x^0 was revealed preferred to x^1 – then consistent behavior would imply that if bundle x^1 was chosen at prices p^1 , it was because x^0 was not affordable at p^1 . In other words:

$$\sum_{i=1}^n p_i^0 x_i^1 \leq \sum_{i=1}^n p_i^0 x_i^0 \rightarrow \sum_{i=1}^n p_i^1 x_i^0 > \sum_{i=1}^n p_i^1 x_i^1. \quad (\text{WARP})$$

The main result of Samuelson 1938a was that in the general n -good case, WARP and the other restrictions imposed on demand functions, implied that the demand functions satisfied two out of the three standard restrictions from ordinal utility theory: negative substitution effects ($S_{ii} < 0$ for all i) and a negative semi-definite Slutsky substitution matrix ($x^T[S]x \leq 0$ for all $x \neq 0$), where

$$S_{ij} = \frac{\partial x_j}{\partial p_i} + x_j \frac{\partial x_i}{\partial M},$$

S is the $n \times n$ matrix with representative element S_{ij} , and T indicates transpose. In the general n -good case WARP did not imply the third restriction from ordinal utility theory – Slutsky symmetry ($S_{ij} = S_{ji}$ for all i and j) – but this was not a problem for Samuelson.² The Slutsky symmetry condition was an integrability condition that guaranteed the existence of a rationalizing utility function,³ and since Samuelson's stated purpose was the elimination of utility from consumer choice theory, the absence of a utility function was not a problem for the new theory. As he said: "I cannot see that it is really an important problem, particularly if we are willing to dispense with the utility concept, and its

² In special case of only two goods ($n = 2$) WARP does imply all three conditions, but not in the general case. The key counterexample was Gale (1960).

³ See Chipman (1976), Chipman and Lenfant (2002), and Hands (2006, 2011a) for historical discussions of integrability, and Hurwicz (1971), Hurwicz and Richter (1979), and Hurwicz and Uzawa (1971) for the technical results. Integrability is discussed in more detail below.

vestigial remnants" (ibid., p. 68).⁴ He was also doubtful that the symmetry (integrability) condition was empirically observable: "I have little faith in any attempts to verify this statistically" (ibid.). Although Samuelson's new consumer choice theory implied two of the traditional restrictions on demand functions, the absence of symmetry meant that his approach was a non-integrable theory of demand: "I should strongly deny, however, that for a rational and consistent individual integrability is implied, except possibly as a matter of circular definition" (ibid.). An undated early draft of Samuelson (1938a) titled "New Foundations for a Pure Theory of Consumer's Behavior" is contained in the Samuelson archives (Box 152) and while the draft is longer and a bit more critical of ordinal utility theory than the published version (note that "New Foundations for" became "A Note on"), the main themes are clearly the same.

Finally, it should be noted that Samuelson's claim that he had a "new" theory that was an improvement on ordinal utility theory is not just a feature of 1938a. He also refers to the methodological importance of his results in a paper on the empirical implications of ordinal utility theory published later in 1938. In that paper he says:

Recently I proposed a new postulational base upon which to construct a theory of consumer's behavior. It was there shown that from this starting point could be erected a theory which included all of the elements of the previous analysis. There I expressed my opinion as to the advantages from a methodological point of view of such a reorientation.

(Samuelson, 1938c, p. 346)

The next major development in the revealed preference program came with the publication of Hendrik Houthakker's "Strong Axiom of Revealed Preference" (SARP) in 1950. Houthakker was at the time fresh out of graduate school at the University of Amsterdam and was working in Richard Stone's applied economics group at Cambridge University; it was his first, and undoubtedly his most influential, paper. The strong axiom basically extended Samuelson's WARP to all finite sequences of price-quantity combinations in such a way that the final bundle in any sequence could not be revealed preferred to the initial bundle. Unlike the weak axiom, SARP implied all three of the standard restrictions on consumer demand functions including Slutsky symmetry (integrability). As Houthakker explained in the paper's first few sentences:

⁴ It is important to note that Samuelson did not use the term "revealed preference" in the original 1938 paper, which certainly makes sense given that the goal was to eliminate utility and preference from consumer choice theory. Samuelson started using the term revealed preference for his new consumer choice theory in Samuelson (1948a), although the term appears briefly in Foundations (1947), but it is in the discussion of index numbers, not in the discussion of consumer choice theory (and this was not changed in the enlarged edition of Foundations in 1983).

Professor Samuelson's "revealed preference" approach has proved to be a useful basis for deriving a considerable part of the static theory of consumer's choice. Its existing versions are not sufficient however, to determine whether or not consumer's preferences can be described by a utility function of the customary type (the problem of integrability), except in the unrealistic case of two commodities. (Houthakker, 1950, p. 159)

The integrability condition meant that for any set of demand functions satisfying SARP and Samuelson's other conditions, there would always exist a rationalizing utility function: a utility function that if maximized subject to a budget constraint, would generate the same demand functions. This also means, of course, that ordinal utility theory and SARP-based revealed preference theory were equivalent: one could start from the consumer maximizing a well-behaved ordinal utility function subject to a budget constraint, or from demand functions satisfying SARP, and the empirical restrictions on demand functions would be exactly the same. As Samuelson explained years later in a letter to Amartya Sen:

For the case $n=2$, the Weak Axiom applied to every binary observation does completely exhaust the content of the general theory. In that sense it is fruitless to argue whether the revealed preference approach is better or worse than earlier versions: ... When n exceeds 2, we must go beyond binary weak axioms to strong axioms. Thanks to Houthakker, we can then show that the generalized revealed-preference formulation completely exhausts the observable P, Q implications of conventional choice theory. So it is equally unimportant to ask whether it is better or worse than those alternatives – like asking whether Newton's Principia is better or worse in English than in Latin. (April 4, 1973, Box 68)

Houthakker's work set off the "high theory" version of revealed preference. The core results from WARP and SARP were expanded and elaborated in various technical ways. How much can the various mathematical assumptions be weakened and still retain the basic results? How about inequalities, boundary conditions, alternative domains? What are the implications of still other revealed preference axioms such as Weak WARP (Kihlstrom, Mas-Colell, and Sonnenschein 1976)? And these were only a few of the numerous questions and concerns addressed in this technical literature.

Although the WARP- and SARP-based high theory literature produced a number of important, and mathematically quite sophisticated, results, it was not very well-suited for applied demand analysis and did not generate a substantial empirical literature.⁵ There were at least three reasons for this. First, the WARP- and SARP-based literature started from demand functions defined over all possible prices and income parameters. In Robert Pollak's (1990) apt language, the high theory literature was "standard domain" revealed preference theory because it started from the same domain as ordinal utility theory: all possible prices and income levels (all budget sets). On the other hand, economists doing empirical work seldom have choice data for all budget sets; they have only quantities chosen for a finite number of price-income combinations and thus require what Pollak calls "restricted domain" theory.⁶ The second problem is that for applied problems it would be useful to have (at least an estimate for) the consumer's utility function. The strong axiom guarantees the existence of a rationalizing utility function, but did not provide any way of constructing or estimating such a function.⁷ Despite the name "revealed preference," the high theory literature did not actually provide the tools to "reveal" preferences or utility. Finally, the one paper that suggested a bridge between revealed preference theory and the actual revelation of preferences, Samuelson (1948a), was not much help. What Samuelson demonstrated in that paper was that in the case of only two goods ($n = 2$) it was possible to construct indifference curves from WARP-consistent consumer choices. Since indifference curves provide the information necessary for an ordinal utility function, Samuelson's paper opened the door from revealed preference to the construction of a utility function, but his results were limited to two dimensions and were not directly useful for empirical applications.

The tools that set the stage for the resolution of this tension between revealed preference theory and empirical (restricted domain) demand analysis came as a result of the publication of Sidney Afriat's important paper in 1967 and the associated theoretical literature (Diewert 1973; Fostel, Scarf, and Todd 2004, etc.).⁸ What Afriat proved in

⁵ There were a few empirical tests of revealed preference theory, but the literature was not extensive. Moscati (2007) provides a detailed discussion of the pre-1970s empirical literature.

⁶ By assuming well-behaved demand functions the high theory literature simply avoided the question of how the relevant demand functions might be obtained from finite choice data. As Samuelson made clear in his original paper, the assumption was that demand functions were potentially "observational," but observational is not the same as actually observed. Revealed preference-based empirical work required starting from actual (finite) observations, not potentially observable demand functions.

⁷ Notice it is always "a" utility function and not "the" utility function. All monotonic transformations of a utility function will generate the same demand functions, so "We cannot hope to find the preference generating choices, but only some preferences – a set of 'equivalent' preferences" (Richter, 1987, p. 167).

⁸ Afriat's paper was one that even influential economic theorists have called "virtually impenetrable" (Pollak, 1990, p. 148), so the secondary literature that filled in the details of how a utility function might

a key theorem was that if a version of the revealed preference axiom – what came to be called the "Generalized Axiom of Revealed Preference" GARP – holds, then a particular set of linear inequalities has a solution. Since these inequalities can be solved for an associated utility function, the link between revealed preference and the construction of a utility function was established. This in turn led to the development of an extensive, and still growing, literature on empirical revealed preference analysis (Bronars 1987, Gross 1995, Varian 1982, 1985, and others); see Varian 2006 and Moscati and Tubaro 2011 for discussions of this research. This empirical literature is also called "nonparametric" demand analysis because unlike the traditional econometric approach to demand analysis it is not necessary to estimate parameters for particular functional forms of the utility function.

The contemporary version of GARP-based demand analysis works in the following way. Starting with a finite set of choice data (prices and corresponding quantities) that has been tested against, and found to be consistent with, a version of GARP – noting that there are various "goodness of fit" tests to relax the conditions for data that is "almost" GARP consistent (Gross 1995, Varian 1985, 2006) – a version of Afriat's theorem is used to construct a particular utility function consistent with (that rationalizes) the data. This utility function is then used for predicting the consumer's response to changes in parameters and characterizing consumer choice in various economic models.

This is a very different kind of revealed preference analysis than that of Samuelson, Houthakker, or the high theory literature. There is of course a domain difference – all possible parameters for the standard domain literature, and finite choice data for the restricted domain GARP-based empirical research – but there are also other differences. The earlier literature started from demand functions and focused on the relationship between the restrictions that revealed preference imposed on those demand functions and those imposed by ordinal utility theory. The GARP-based empirical research focuses much more on empirical application. It starts with finite choice data and then uses consistent patterns found in that data, and a corresponding utility function, to project similar patterns onto other sets of parameters. Both involve consistency conditions, but the more traditional revealed preference literature focuses on qualitative comparative statics (Slutsky symmetry, etc.) and the mere existence of a rationalizing utility function (in the SARP case); it starts from functions assumed to be observational, but is aimed more at general theory than specific empirical applications. The empirical revealed preference literature starts from finite choice data, constructs a particular utility function that (approximately) fits the data, and then uses the information in specific empirical applications; it is as

actually be constructed played an extremely important role in carrying out the empirical program suggested by Afriat's paper. See Varian (2006) for a detailed discussion.

much a technique for empirical inference – competing with econometrics – as a theory of consumer choice. They are, as I have been saying, members of the same family, but they are also quite distinct members of that family.⁹

The final member of the revealed preference family to discuss is the most recent version, what I call "contemporary revealed preference theory" (Hands 2011b, 2011c). This approach starts from the general framework of GARP-based empirical demand analysis – begin with finite choice data and use consistent patterns in that data to construct, via a version of Afriat's theorem, a utility function which can be used to project the consumer's behavior onto other data sets – but then gives it a methodological twist. Defenders of contemporary revealed preference theory – and I will use Kenneth Binmore (2009a, 2009b) and Faruk Gul and Wolfgang Pesendorfer (2008) as paradigmatic examples¹⁰ – argue that revealed preference theory exhausts all of (at least choice theory in) positive economics: it is the "official doctrine of neoclassical economics, enshrined in all respectable textbooks" (Binmore, 2009a, p.20).¹¹ Not only is it what (all that) economist actually do in choice theory, it is all they should do, because it is the only theoretical activity supported by the type of data available to economists. As Gul and Pesendorfer explain:

The relevant data are revealed preference data, that is, consumption choice given the individual's constraints. These data are used to calibrate the model (i.e. to identify particular parameters), and the resulting calibrated models are used to predict future choices and perhaps equilibrium variables such as prices. Hence, standard (positive) theory identifies choice parameters from past behavior and relates these parameters to future behavior and equilibrium variables. (Gul and Pesendorfer, 2008, pp. 7-8)

⁹ It should be noted that there has been a substantial methodological literature on revealed preference theory – both praising and criticizing it – over the years: See Georgescu-Roegen (1954), Gul and Pesendorfer (2008), Hausman (2000, 2008), Little (1949), Rosenberg (1992), Sen (1973, 1993, 1997) and Wong (2006) for a sample. As I have argued elsewhere (Hands 2011b, 2011c) once one understands the differences between the various versions of revealed preference theory it becomes clear why such methodological analysis often seems so ineffective. The problem is that that given the differences among the various versions of revealed preference, a methodological analysis aimed at one version is unlikely to say much, either way, about any of the others.

¹⁰ I will also quote from Douglas Bernheim and Antonio Rangel (2008) because they provide particularly clear statements of the position. I did not include them among the paradigmatic examples because it is not clear whether they would support the strong methodological claims I identify with contemporary revealed preference theory.

¹¹ They also extend the argument to welfare economics, although I will restrict the discussion here to positive consumer choice theory. See Hands (2011c) for a discussion of the issues surrounding welfare economics.

There are many features differentiating this version of revealed preference theory from other members of the family in addition to its methodological tone, but I will note only two (and they are closely related). The first is the complete identification of "preferred" with "chosen," and the second is that preferences and/or utility do not explain, and are not causally connected to, the behavior of the consumers in question. The traditional idea – that preferences cause choices – is what Binmore calls the "Causal Utility Fallacy," and he argues that economists need to give it up: to "give up any pretension to be offering a causal explanation of ... choice behavior in favor of an account that is merely a description of the choice behavior of someone who chooses consistently" (Binmore, 2009a, p. 20). For most applications of rational choice in economics, including ordinal utility theory, the relevant theory is a formalization of the common sense or folk-psychological idea that people have preferences over various states of the world and they act so as to bring about states of the world which most satisfy those preferences. People have preferences and those preferences, along with beliefs and the constraints the individual faces, cause the observed choice behavior. This is not the case for contemporary revealed preference theory, in fact it is just the reverse. According to this most recent version of revealed preference, a bundle is not chosen because it is preferred; it is preferred because it was chosen. As Bernheim and Rangel explain:

Though we often speak of choice as determined from preferences, the opposite is actually the case. Standard economics makes no assumption about how choices are actually made; preferences are merely constructs that summarize choice. (Bernheim and Rangel, 2008, p. 158)

Similarly, Gul and Pesendorfer:

In the standard approach, the terms "utility maximization" and "choice" are synonymous ... Standard economics focuses on revealed preference because economic data come in this form ... Such data do not enable the economist to distinguish between what the agent intended to choose and what she ended up choosing, what she chose and what she ought to have chosen. (Gul and Pesendorfer, 2008, p. 8)

Obviously in both of these quotations "standard economics" means "standard economics as seen by contemporary revealed preference theory." And finally Binmore's Causal Utility Fallacy:

In revealed-preference theory it isn't true that Pandora chooses b rather than a because the utility of b exceeds the

utility of a. This is the Causal Utility Fallacy. It isn't even true that Pandora chooses b rather than a because she prefers b to a. On the contrary, it is because Pandora chooses b rather than a that we say that Pandora prefers b to a, and assign b a larger utility. (Binmore, 2009a, p. 19)

It is important to note that this is a significant change from other approaches to revealed preference. Although certain authors interpret "preference" as just behavior and see revealed preference theory as simply a way of redescribing behavior, rather than a way of explaining it or trying to understand its causes, before contemporary revealed preference theory it was certainly not necessary to interpret revealed preference in this way. The fact that ordinal utility theory has traditionally been viewed as an explanatory theory of individual choice behavior – people have preferences (subjective mental states) that come before, and when combined with the relevant constraints, cause and explain the choices the consumer makes – combined with the fact that (post-SARP) revealed preference is equivalent to ordinal utility theory, implies that it is possible to interpret revealed preference in this way as well. Certainly the language of "revealed" preferences suggests something existing prior to the choice it explains and causes (When the x-ray "reveals" the broken bone, we assume both that the break was there before – and is not synonymous with – the x-ray, and that it explains, and is the cause of, the pain). So too for the GARP-based empirical work. The most natural interpretation of the utility function obtained by these empirical techniques is to think of it like we normally think of estimates in economics – as attempts to measure the relevant causal factors. Now as we will see below there is quite a bit of ambiguity on these issues – even with Samuelson himself – but my point here is not to take sides in the debate about whether contemporary revealed preference theory is methodologically a better or worse interpretation of revealed preference theory. My point is simply to note that for previous revealed preference approaches it was possible to think of preferences as separate from the choices they cause and explain. In the case of contemporary revealed preference theory this is no longer an option and that makes it a quite different from other members of the revealed preference family.

2. Das Paul Samuelson and Related Problems

Even the reader who has not previously thought much about revealed preference may have noticed the stark difference between Samuelson's statements in the 1938 paper where the goal was to eliminate the last vestiges of utility theory, and his post-SARP position where revealed preference was just a different characterization of ordinal utility theory. It is indeed a stark difference, and while it is not the only question that needs to be addressed in this section, it is a good place to

start since it is the most common entry point for those who argue that Samuelson was inconsistent in his views over time (Das Paul Samuelson Problem).

The published text that seems to make the sharpest contrast with the "eliminating the last vestiges of utility" language or the original paper is Samuelson's 1950 paper on integrability. The 1950 paper begins with a history of the integrability question, then makes a number of arguments for why integrability (in two, three, and higher dimensions) is a reasonable assumption in consumer choice theory – why various arguments for non-integrability are "red herrings" (Samuelson, 1950, p. 360) – and then closes with a mathematical addendum which proves that Slutsky symmetry is necessary and sufficient for integrability. At one point in the second half of the paper Samuelson summarizes how he sees the path that brought him to revealed preference and its relationship to integrability:

We are now in a position to complete the programme begun a dozen years ago of arriving at the full empirical implications for demand behaviour of the most general ordinal utility analysis. My own work in this direction grew out of a remark made to me by Professor Haberler in his 1936 international trade seminar at Harvard ... I suddenly realized that we could dispense with almost all notions of utility: starting from a few logical axioms of demand consistency, I could derive the whole of the valid utility analysis as corollaries ...

I soon realized that this could carry us almost all the way along the path of providing new foundations for utility theory. But not quite all the way. The problem of integrability, it soon became obvious, could not yield to this weak axiom alone. I held up publication on the non-contradictions of revealed preference for a chain of three or more situations, then non-integrability could indeed be excluded. At scientific meetings and in correspondence this problem was proposed, both to economists and mathematicians.¹² But no proof was forthcoming for all these years, until Mr. Houthakker's paper arrived in the daily mail.
(Samuelson, 1950, pp. 369-70)

This certainly sounds like a capitulation on the goals that were so clearly stated in the 1938 paper: removing preference and utility from consumer choice theory, and developing a non-integrable theory of

¹² In several writings, published and correspondence, Samuelson names the economists and mathematicians he consulted. He mentions Stan Ulam and Lynn Loomis in Samuelson (1998, p. 1380) and Ulam and Rollin Bennett in letters to Houthakker (July 12, 1961, Box 38) and Marcel Richter (April 20, 1964, Box 62).

demand. Not only does it sound like Samuelson changed his mind on these matters, the story in 1950 – and the story Samuelson consistently told thereafter – is that he was always motivated by finding an equivalent (and thus integrable) theory: one with the same "full empirical implications" as ordinal utility theory. Notice that in the first sentence he even refers to "the most general ordinal utility analysis," which clearly reflects a change from his earlier attempts to develop a non-integrable theory of consumer choice. In the 1930s debate over integrable and non-integrable demand, those like R. G. D. Allen and Griffith Evans who emphasized non-integrability, viewed it as the general case – the existence of a utility function was a special case, non-integrability subsumed that special case, but also allowed for demand theory in the absence of a well-behaved utility function.¹³ By 1950 Samuelson seems to have accepted the view that ultimately became the standard story – that consumers having a well-behaved utility function was the "general" case and non-integrability was an unimportant special case. This is far different from his 1938 position. Samuelson's 1950 statements on utility and integrability clearly seem to be inconsistent with the view he stated earlier. Of course it is not inconsistent to change one's mind, but it is inconsistent to say something entirely different at a later date and then insist that the later view was the one you always held. The statement of the program's goals were entirely different in 1938 than in 1950, which suggests that either Samuelson changed his mind about utility theory and never admitted it, or that he did not really believe what he wrote in 1938 (and given the vociferousness of the language in that paper and in the early draft, the latter seems very unlikely). In any case, it clearly raises the question of Das Paul Samuelson Problem. The next section will speculate a bit about what may have motivated Samuelson's statements, but at this point it seems useful to continue on with what Samuelson and others had to say about Samuelson's relationship to revealed preference theory.

Samuelson's argument that he was, from the start, seeking a theory with the "full empirical implications for demand behaviour of the most general ordinal utility analysis" including integrability, and that he initially had the basic idea of the strong axiom but could not provide the proof, is repeated in various forms throughout the rest of his life. For example, his very first letter to Houthakker starts out:

Many thanks for sending me your paper on "Revealed Preference and the Utility Function." It is a gem. It completely answers my conjecture of a decade ago – and in the affirmative. Back in the 1930s I asked a number of mathematicians and mathematical economists to help me prove that if my first weak axiom on non-contradiction of

¹³ See Hands (2011a and 2006) for a discussion of Allen's and Evan's views on these matters.

revealed preferences on pairs of commodities were strengthened to a second axiom on non-contradiction for any chain of situations, then non-integrability would be ruled out. A number of people replied that this seemed to them reasonable, but no one supplied the proof and I was not able to devise one. (December 23, 1949, Box 38)

Similarly, in a letter to John Hicks dated January 25, 1952, Samuelson says that he conjectured SARP but could not prove it, and that he has "always regarded revealed preference as a very unpretentious reformulation of ordinal demand theory" (Box 37). In a later letter to Houthakker dated July 12, 1961, he says "Already in 1937 I realized that when you went beyond the two good case the weak axiom was not strong enough" and "I have never felt that the symmetry of the Slutsky coefficients lacks importance" (Box 38).¹⁴ And in a letter to Marcel Richter April 20, 1964, he explains that early on he thought that SARP was sufficient for integrability but he had a "failure of nerve" about publishing the results. And such statements are not just in correspondence, they appear in published work as well. For example, similar remarks were made in his Nobel lecture:

From the beginning I was concerned to find out what refutable hypotheses on the observable facts on price and quantity demanded were implied by the assumption that the consumer spends his limited income at given prices in order to maximize his ordinal utility ... To make a long story short, the flash of inspiration for "Revealed Preference" came to me in argument with one of my teachers, as so many of my best ideas have been. (Samuelson, 1972, p. 256)

This is particularly telling since his Nobel lecture was on "maximum principles" and revealed preference was initially supposed to replace maximization with consistency. He also repeated the story in detail in "How Foundations Came to Be" in 1998 as well in various other biographical remarks.

The inconsistency associated with Samuelson's goal of eliminating utility from consumer choice theory and then turning revealed preference into a tool for reinforcing that same theory, has certainly not gone unnoticed within the literature. Even Houthakker, a long supporter of revealed preference and someone with a long and close professional

¹⁴ It is interesting that in this letter he says "I write for you only" and "for your ears alone" and yet what he is saying is much like what he said in Samuelson (1950).

relationship with Samuelson,¹⁵ raised the question in a 1983 volume on Samuelson's many contributions to economics:

A final word about the significance of the revealed preference approach. Did it succeed in "develop[ing] the theory of consumer's behavior freed from any vestigial traces of the utility concept," to quote the closing paragraph of Samuelson (1938a ...)? Or did something happen on the way? One is tempted to conclude that something did happen, for in 1950 Samuelson restated "the program begun a dozen years ago" as one of "arriving at the full empirical implications for demand behavior of the most general ordinal utility analysis" ... The stone the builder rejected in 1938 seemed to have become the cornerstone in 1950. (Houthakker, 1983, p. 63)

Although a few sentences later Houthakker notes that compared to Samuelson's overall achievement with revealed preference theory, his "animadversions about utility theory are of secondary importance" (ibid.), he still felt it necessary to note the tension. Pollak similarly acknowledges the issue in his discussion of Houthakker's contribution – "Although Samuelson's 1938 article flirted with the possibility of divorcing consumer theory from notions of preference and utility, the principle focus of revealed preference theory soon shifted to establishing the equivalence of axioms about preference and axioms about demand" (1990, p. 142).

There are many other examples in the literature, but the most sustained defender of the inconsistency thesis was Stanley Wong in Foundations of Paul Samuelson's Revealed Preference Theory (2006, originally published in 1978). Wong argued that not only did Samuelson exhibit an inconsistency on this matter – and inconsistency between the "last vestiges" language of 1938 and the "full implications of ordinal utility theory" in 1950, but also between his 1948 paper on constructing indifference curves and the other two papers. Wong argued that Samuelson had three separate goals for revealed preference theory (or as Wong put it "problem situations"): i) to eliminate utility from consumer choice theory in 1938, ii) to use revealed preference to actually construct indifference curves in 1948a, and iii) to deduce the full empirical implications of ordinal utility theory by other means in 1950. Thinking back to the previous section, the degree to which each of these goals manifests itself within the revealed preference literature depends substantially on which version of revealed preference one is focusing on.

¹⁵ The Samuelson archives contain an extensive correspondence between Samuelson and Houthakker that contains over sixty letters, starting in December 23, 1949 with Samuelson thanking Houthakker for sending him a pre-publication copy of the SARP paper and ending with Samuelson's May 7, 2008 letter to the family expressing condolences for "the sad news of Hendricks's passing" [Box 38].

The high theory literature is most indebted to the third of these goals and is most directly linked to integrability and Samuelson (1950), while the GARP-based empirical work is most directly descended from – though its empirical practicality was much improved by later developments – Samuelson (1948a). Wong is undoubtedly correct as a matter of logical analysis; there are inconsistencies, and they are inconsistencies that Samuelson never admitted in his writings or explained in any satisfactory way. The effectiveness of Wong's criticism was also enhanced by the particular approach that he used to criticize Samuelson: basically applying revealed preference to Samuelson. At the heart of all versions of revealed preference is the idea that rationality is equivalent to consistency, and what Wong demonstrated was that Samuelson himself was inconsistent over time, and thus, by implication, irrational in his theorizing about consumer choice.

Samuelson's response to Wong will be discussed briefly in the next section, but here I would only like to make two points about Wong's criticisms. First, although the inconsistencies Wong identifies are there in Samuelson's writings – and in some sense replicated in his correspondence which was not available until quite recently – I would say the main inconsistency is between what he said in the 1938 and what he said everywhere else. If the 1938 paper had never been published, Samuelson's view of revealed preference would seem pretty consistent over time. There are differences between the 1948 paper and the 1950 paper, but they seem to be the same type of differences that exist between the high theory literature and GARP-based empirical demand analysis; they are differences between pure theory and applied empirical analysis that exist in every area of economics. The second point is simply that even if Wong is entirely correct about all three inconsistencies, it is primarily a philosophical issue; my project here is try to get a better historical understanding of why Samuelson – surely a gifted mathematician with an eye for inconsistency – wrote what he wrote.

One approach to reconciling some of the tension within Samuelson's writings on revealed preference would be to look to Foundations (1947), since it was published almost a decade after the original paper (and after his years in the Harvard Society of Fellows), was in fact a mathematical "foundation" for all of economic analysis, and was republished in an enlarged edition in 1983. One might suspect that if there had been a change during the 1938-1947 period it would have shown up in Foundations, and if not there, then later in the enlarged edition. Unfortunately Foundations only adds to the confusion about the presence or absence of a change of mind.

The problem with Foundations – both the original and the later enlarged edition – is that revealed preference plays almost no role. The original 1938 paper is cited (p. 111) in chapter five on consumer choice theory, but the entire chapter is presented in terms of what is now the standard ordinal utility theory point of view (Lagrangian functions, first

and second order conditions, etc.). The brief mention of Samuelson (1938a) is to point out another way of characterizing the negative substitution effect. He notes that "Elsewhere I have suggested as a new foundations for the pure theory of consumer's behavior" i.e. the WARP condition, and that "this single condition provides us with complete foundations for the theory (with the reservation concerning integrability)" (ibid.), but the discussion is only a few pages. The consumer choice chapter in Foundations is an ordinal utility chapter, with only a passing reference to revealed preference. The term "revealed preference" does appear in Foundations, but in the discussion of index numbers, not in consumer choice theory, and in a section where he assumes ordinal utility theory.¹⁶

Another thing that adds to the confusion about how Samuelson presented revealed preference theory in Foundations, is that although he chose to present consumer choice in ordinal utility terms, he was generally quite pessimistic about the empirical usefulness of the theory he presented. An excellent example is the first epigraph above, but there are other examples scattered about in chapter five (and, by the way, chapter five was reproduced without change in the enlarged 1983 edition). For example:

I do not propose to defend the fruitfulness of these empirical restrictions. The extent to which they satisfy and unify the factual behavior of consumer cannot be settled by argumentation. However, for better or worse the theory of utility has occupied an important position in economic thought for the last half century. This alone makes it desirable that its meaning be clearly understood.
(Samuelson, 1947, p. 92)

These are not very positive words. Samuelson's tone was so negative on consumer choice theory that it was noticed by several reviewers. William Baumol (1949) noted that "Samuelson argues that the content of the received theory of consumer behavior is rather limited" (p. 163); Lloyd Metzler noted his "somewhat cynical" attitude about the theory of consumer choice (p. 906); and George Stigler said that for Samuelson, empirical contradiction "would not matter much in the case of utility theory" (p. 117). Even though these were just book reviews, both Baumol and Metzler evidently thought the point was important enough that they quoted Samuelson's remarks on page 117 of Foundations (the first epigraph). If Samuelson thought ordinal utility theory was so problematic, then why didn't he make chapter five into a revealed preference chapter? And even more interesting, why did he say only a few years later that revealed preference was nothing more than another

¹⁶ It seems this was first pointed out by Wong in 1978 (2006, p. 67).

version of (by 1950 evidently just fine) ordinal utility theory? As I said, Foundations only raises more questions.

Another question that might be raised about Samuelson's own view of revealed preference theory concerns description versus (or and) explanation. Did Samuelson see revealed preference theory as a way of explaining consumer choice – as has traditionally been the case with ordinal utility theory – or would he see it as only a tool for economical description (or redescription) as is the case for contemporary revealed preference theory? Again, it is simply not clear from what Samuelson himself wrote in either published work or in correspondence.

On one hand, in Samuelson's few explicitly methodological writings (Samuelson 1963, 1964, 1965) he takes a narrowly empiricist position that theories describe only empirical phenomena and do not explain them (a position he traces to the early positivists that he calls "empirical realism"¹⁷). He presents this position quite clearly in his 1964 reply to Fritz Machlup:

Scientists never "explain" any behavior, by theory or by any other hook. Every description that is superseded by a "deeper explanation" turns out upon careful examination to have been replaced by still another description, albeit possibly a more useful description that covers and illuminates a wider area. (1964, p. 737)

And again, even more strongly a year later:

[T]here is a widespread will to disbelieve in my rather hard-boiled insistence upon "theory" as (strategically simplified) description of observable and refutable empirical regularities, and a widespread hankering for a more exalted ... role of theoretical explanation ... An explanation, as used legitimately in science, is a better kind of description and not something that goes ultimately beyond description. (1965, pp. 1164-65)

And what does he cite as one of the best examples of such successful descriptive, but non-explanatory, science? Revealed preference theory of course:

The doctrines of revealed preference provide the most literal example of a theory that has been stripped down to its bare implications for empirical realism: Occam's Razor has cut

¹⁷ This, and Samuelson's other use of the term "realism" in methodological discussion, takes some getting used to since the philosophical position he endorses would, in contemporary philosophical parlance, be an instrumentalist, not a realist, position.

away every zipper, collar, shift, and fig leaf. (Samuelson, 1964, p. 738)

But he also, in numerous places – and by the way, philosophically consistent with his empiricist descriptivism – says that if two theories have exactly the same empirical content, then the two theories are identical. In particular he notes the following relationship between two theories A and B, and C a set of empirical consequences:

If C is the complete set of consequences of B, it is identical with B ... There can be no factual correctness of C so defined that is not also enjoyed by B. The minimal set of assumptions that give rise to B are identical with B, and if A is given this interpretation, its realism cannot differ from that of the theory B and consequence C. (Samuelson, 1963, p. 234)

In other words, ordinal utility theory and (strong) revealed preference theory have the same exact consequences and are therefore descriptively – as Samuelson would say "realistically" – equivalent. This of course implies that ordinal utility theory is also purely descriptive and non-explanatory. This argument has been the basis for much of the philosophical criticism of revealed preference theory over the years: that it is simply a descriptive, and not an explanatory theory (Hausman 2000, 2008; Rosenberg 1992; Wong 2006 and others). Perhaps this is a valid methodological criticism, but our concern here is with historical issues, and such descriptivism flies in the face of what the majority of economists think and say about ordinal utility theory (traditionally and today),¹⁸ but it is also, and more importantly here, in conflict with what Samuelson himself sometimes said. In the consumer choice chapter of Foundations he discusses utility theory and turns to "the use to which these concepts are put in the explanation of consumer's behavior" (1947, p. 96), and in the 1950 integrability paper he continues to disconnect choice theory from welfare economics by arguing that "if integrability should turn out to be the hypothesis to explain the empirical facts of the market-place, this makes a belief in individualistic ethics possible but still not mandatory" (p. 375). Maybe these, and similar remarks

¹⁸ It is obvious to me this is the case and I suspect that is the case for most readers, but it is also an empirical claim about what most economists have (and do), thought and said (think and say), and thus something that is notoriously difficult to defend. For the most part I will just assume this is as obvious to the reader as it is to me, but I cannot resist a two quick quotes supporting this claim. The first is from Marcel Richter's New Palgrave entry on revealed preference: "From the economist's point of view, unobservable preferences generate observable choices" (1987, p. 167). The second is from Binmore who is defending (non-explanatory) contemporary revealed preference theory by arguing that "we have to give up any pretension to be offering a causal explanation" of choice behavior. If economists need to give it up, it would seem to be the view they currently hold.

elsewhere, are just places where he is not being sufficiently careful, but in any case we are still left with questions about why he never made a clear statement on the descriptive and/or explanatory status – either methodologically or historically – of ordinal utility theory, and how it related to (post-SARP) revealed preference theory: a statement that would have eliminated much of the existing ambiguity.

In closing this section there are a few more minor issues – not a contradictions, but things that contribute to the existing lack of clarity about how Samuelson interpreted revealed preference theory. One problem is that it does not appear that Samuelson ever commented in any serious way on the GARP-based empirical literature.¹⁹ There are a few letters between Samuelson and Afriat (Box 9), but they do not discuss serious questions in either theory or methodology (they are primarily associated with Afriat's efforts to obtain a position in an American university); there is also some correspondence with Hal Varian (Box 74) – who had been an undergraduate at MIT – but there is no discussion of Varian's work on nonparametric demand analysis. Given that Samuelson was concerned from the beginning with the "observational" character of revealed preference theory and its finite (i.e. non-calculus) mathematics²⁰, and given that the GARP-based empirical demand analysis was starting to appear regularly in economics journals after the mid-1980s (and almost always referenced Samuelson 1938a), one would think he would have had something to say about whether this literature was a stride forward, a wrong turn, or whatever. It may just be the result of Samuelson's lack of interest in such empirical research, but again, some commentary would certainly help clear up a lot of questions. By the way, it seems the same can be said for the abstract choice function approach to revealed preference theory (extended domain theory). He did respond to Sen about some of Sen's revealed preference work of this genre (Box 68), but there was not any indication of whether Samuelson thought the general choice function approach was an important step forward, or too abstract, or an entirely different theory.

The bottom line on this section is that Samuelson does not provide much help on the task of understanding exactly how he interpreted revealed preference theory or which of the various later sub-programs he most supported. There seems to be a Das Paul Samuelson problem, at least between his first revealed preference paper and everything he wrote thereafter, but his writings – both published and unpublished – fail to provide any significant insight into how he himself viewed the tension. Samuelson was the father of revealed preference theory, but he did little to reveal his preferences about it.

¹⁹ I have not been able to examine all of Samuelson's unpublished papers and correspondence – so something may yet turn up – but I have looked in all of the obvious places.

²⁰ This emphasis is particularly clear in Samuelson's introduction to the enlarged edition of Foundations.

3. Paul Samuelson and Revealed Preference Theory

In this section I will try to bring the various pieces together in a way that gives us a better understanding of Samuelson's relationship to revealed preference theory. It will be a three step process. First, I will start with some discussion of Samuelson's own response to the charge of inconsistency. Then I will lay out what seems to be the most defensible rational reconstruction of what Samuelson wrote on these issues. Finally, I will discuss some of the open historical questions and why such a rational reconstruction should not be the end of the historical inquiry.

One very important point that has not, thus far, been addressed is that Samuelson seemed to really dislike the appearance of inconsistency. This is very clear from both published and unpublished writing. Samuelson demonstrated this as early as 1948 in a letter to I. M. D. Little commenting on a draft of the paper that became Little (1949):

In my 1938 article, referred to in the book, the emphasis is slightly different than in the book ... You might care to indicate this at certain places I have marked. Probably few people have read the article, and certainly the originality of your own contribution is unaffected. But I should not like for those few who have read it to receive the impression from your account that I have altered my own earlier views. (Samuelson to Little, June 21, 1948, Box 48, emphasis added)

Two things to notice here: the first is that Samuelson, even as early as 1948, was concerned that readers would think he changed his mind about revealed preference, and second that he focused his attention on perceived differences between the 1938 paper and Foundations.²¹ Given that the paper that contrasts most starkly with 1938a does not appear until the November 1950 issue of Economica it is possible it was not written at this point (although even later, well after the publication of the 1950 paper, he continues to express concern that Foundations might be interpreted as a change on the subject of revealed preference).

In a letter to John R. Hicks four years later he expands on his concern over being viewed as changing his mind, and introduces the idea that it was his humility that led him to downplay revealed preference in Foundations:

²¹ In a footnote on page 95, Little (1949) noted that in 1938a "Samuelson takes a more austere attitude towards assuming the existence of indifference curves," and on page 96 that "in his article in Economica, 1938, the treatment is more nearly akin to the present formulation," but it is not clear whether these notes were in the draft that Samuelson commented on, or whether they were added in response to Samuelson's letter.

That I have leaned over backward in differentiation of revealed preference from demand theory generally is perhaps shown by the fact that I contented myself with writing a brief 1938 note. In Foundations I unobtrusively gave its main results without emphasizing its novelty or superiority; and I apparently succeeded so well in this that a number of people have claimed that I backed down from my 1938 position – which I am not aware of having done. (Samuelson to Hicks, January 25, 1952, Box 37)

Notice here that the concern is still the perceived tension between the 1938 paper and Foundations, even though this is post-SARP. Also note that the idea is creeping in that he is being punished in some way for not wanting to over-inflate the differences between his new theory and ordinal utility theory. This of course still leaves the question of why he had the strong words in 1938 and why he is so concerned with the possible reaction.

One of his clearest statements is in an (unfortunately) undated manuscript titled "The Revealed Preference Approach: A Digression." His use of the expression "through this book" would suggest it was a draft of the introduction to the enlarged edition of Foundations, but if so, it did not make it into the printed version.

In 1938, I proposed a new logical foundation for the theory of consumer's choice. Because I tucked its results and philosophy unobtrusively through this book, a number of postwar writers have believed that my position changed between 1938 and 1946, which is not the case. Once I perceived its full logical equivalence with the modern theory of ordinal preference, I felt the first task of squeezing out of that theory its full empirical implications and testable hypotheses had been completed and that there was no need to play up the merits or novelty of the new formulation. (Box 133)

It is clear that Samuelson did not like the idea of being accused of changing his mind, but the rest of the argument is less clear. Foundations appeared in 1947 long before he received Houthakker's draft in 1949. It seems unlikely that perceiving "the full logical equivalence" influenced the content of Foundations since there was no proof of the equivalence at the time he was making decisions about the content of Foundations. There is also the issue that even long after the 1950s Samuelson does not see the main tension between the original 1938 paper and the 1950 paper on integrability, but rather that he downplayed revealed preference in Foundations.

One of the places where one would expect to find Samuelson's most direct response (or defense, or at least clarification) would be in his correspondence with Wong. There are a few letters between the two between 1973 and 1975 (Box 78). The correspondence suggests that they talked at some point in London, Ontario, and Wong recounts some fairly strong statements by Samuelson, including "You retorted that this [Wong's argument] is incorrect and cited as evidence the fact that you have never been interested in the so-called non-integrable case" and "You indicated that you would publish a retraction if I were right and I think that this quotation confirms my view" (Wong to Samuelson, October 14, 1975, Box 78). Unfortunately Samuelson's response to these claims is not available, and his later published remarks on Wong seem to be decidedly non-clarifying:

The deep points raised by Dr. Wong can, I believe, be argued out in the 2-good case without prejudice to their evaluations. Presence or lack of presence of Giovanni Antonelli's ... observable integrability conditions introduce interesting technical points but Wong's preoccupations will remain to be addressed even when integrability is assured (as in the 2-good case). (Samuelson, 1998, p. 1981)

One letter showing that Samuelson's concern over being accused of inconsistency and his unwillingness to explain the tension that so many others see extends into the 1990s, is Samuelson's response to Sen's (direct) question about his "vestigial traces" line in 1938. Sen says:

I personally learned most of the "basics" from the Foundations (like countless others). But I did think on the basis of reading your 1938 paper including your introductory remarks (page 61) and your concluding paragraph (page 71) that you were there developing a "theory of consumer's behavior freed from any vestigial traces of the utility concept," It clearly was read as such by Ian Little, John Hicks, and others. (Sen to Samuelson, March 19, 1992, Box 68, emphasis in original)

Samuelson's response is, at best, evasive. He addressed other issues Sen raised in the body of the letter, but not the "last vestiges" question; on this he added the following post script:

I have not taken the time to discuss my 1938 words "last vestiges of utility," which do not sustain your first-draft innuendo (a non-pejorative word). When I looked up Little, I recognized no flagrant misinterpretations of 1938 Samuelson. If opportunity cost of time were less, as a legal

mercenary I might undertake defense on a contingency basis of Little et al. and Arrow against the stated indictments. Obviously, I cannot be sure of this on the basis of a truncated reading of your text. So take it with a grain of salt. (Samuelson to Sen, April 4, 1992, Box 68)

If what Little said was fine, why did Samuelson ask him to change the original wording? And why, over fifty years, has there never been time to seriously discuss his "last vestiges of utility" remarks? It would be possible to continue on with quotations here, but the spirit of Samuelson's response on these matters should be clear. He was not happy about being accused of changing his mind, and yet he never really explained the relationship between his original paper and his later work on revealed preference (or any number of related issues). Perhaps it is time to move away from Samuelson's actual words and engage in some creative reconstruction.

So how might Samuelson's writings on revealed preference theory be rationally reconstructed? In other words, if one wanted to present the most coherent narrative that can be constructed to rationalize the evidence in his various writings, published and unpublished, how should the story be told? Let me attempt such a reconstruction.

The story begins with Samuelson in 1936-37. The context is his escape from Chicago, his graduate work at Harvard, and the beginning of his period in the Harvard Society of Fellows. These are the heady days of positivism, operationalism, and behaviorism – ideas that Samuelson accepted at the time and never abandoned in explicit discussions of methodological issues (actual practice is a separate issue) – and he is generally skeptical of the utility theory that had so captured the hearts and minds of his Chicago teachers. There is also much discussion in the air about the possibility of a more general version of choice theory: a non-integrable theory that did not start with the assumption that the consumer possessed stable well-ordered preferences defined over the entire choice space. In one of the places where Samuelson came closest to actually explaining what he wrote on revealed preference theory, he identified the influence of Allen, Evans, and others working on non-integrable demand theory in the early 1930s.

Prior to Houthakker, for $n > 2$, I tended to side with Roy G. D. Allen rather than with Hicks's insistence upon integrability. Why not be general and be happy to posit non-integrability and global non-transitivity? In these cases, only the Weak Axiom could be validly posited as a constraint on empirical demand observations. My reading of Griffith Evans (1930), Allen (1932), and Nicholas Georgescu-Roegen (1936) softened me up for such a half-way house compromise. (Samuelson. 1998, p. 1381)

WARP-based revealed preference theory was thus a more general demand theory which predicted some empirical restrictions on demand functions, but moved the profession sharply away from any talk about utility and the associated intentional concepts that were in the process of being purged from psychology (and it did so without being "just" behaviorism²²) while also retaining some of the existing theoretical framework (optimization and the associated mathematical tools, but also the idea that, at least compensated, demand curves sloped downward). Then, in 1949, came Houthakker's paper and it became clear how little was required to make revealed preference and ordinal utility theory observationally equivalent. Of course given Samuelson's operationalist and positivist commitments, observationally equivalent meant "the same theory." As Samuelson put it in the paper quoted immediately above:

When members of a class are equal, each is first among equals, and is last among equals. Why debate the different merits of essential equals? (ibid., pp. 1380-81)

But not only are they equals, given Samuelson's stated methodological preferences, the observational equivalence of the two theories elevates the scientific status of ordinal utility theory. His early criticism of ordinal utility theory was that it was empirically empty – and thus, in his positivist vision, meaningless – and Houthakker demonstrated that was not the case. Revealed preference thus bolstered Samuelson's faith in ordinal utility theory. From that point on his responses become more a story about his personal interests – for example his lack of interest in applied demand theory explains his failure to address GARP-based empirical research, and his lack of interest in Bourbakian mathematical economics explains his failure to address the abstract choice function-based revealed preference literature. For readers who are seeking a neat and tidy narrative – a coherent rational reconstruction – I humbly submit, this is the best available. There is a Das Paul Samuelson Problem in the sense that he changed his mind, but he changed it for what, given his stated methodological position, seem to be fairly good reasons.

This seems to be a reasonable story that makes Samuelson willing to learn from additional evidence (at least mathematical evidence) at any point in time, and yet consistent in his epistemic vision over time. I think this story is more satisfying than simply pointing out inconsistencies in various texts and leaving it at that (although I understand that philosophical and historical research have different goals). I also think it is a much more defensible narrative than the one that I myself had long

²² I have discussed some of the issues associated with behaviorism in economics during this period in Hands (2009).

entertained. My original story was that it was all about the neoclassical synthesis. Proposing a radical new consumer choice theory and arguing for the elimination of utility theory was fine for a brash young mathematical economist, but Samuelson of the neoclassical synthesis – Samuelson the great harmonizer (Pearce and Hoover 1995) – wanted to find common ground between neoclassical economics and Keynesian macroeconomics, and being seen as someone trying to overthrow utility theory – the heart of neoclassicism if it has one – was not conducive to the successful promotion of the neoclassical synthesis. Hence, or so I thought, this was the reason for Samuelson's change in attitude about ordinal utility theory during the mid 1940s. The problem is, I could find very little evidence for it in Samuelson's published work, and essentially none in the archives.

So why not stop here with a reasonable rational reconstruction? The reason is that the (actually any) rational reconstruction leaves out so much of the story. There are still many unanswered questions – questions that may in fact be answered satisfactorily with additional research (in published or archival materials, or the archives of his colleagues and correspondents). I will close this section by just noting a few of these questions.

The first is Samuelson's resistance to being seen as inconsistent and/or changing his mind. In the contemporary parlance "What was that about"? A key question seems to be whether this was a general characteristic of Samuelson's professional work, or whether it was unique to revealed preference theory. There is also the "why comment at all?" question. Samuelson very seldom sheds any serious light on any of relevant issues, and yet he seems continually compelled to say something. As Samuelson himself once said about another economist (Bertil Ohlin): he "repeatedly comes back to the point only to leave it elusively hanging in the air ... This is hardly cricket" (Samuelson, 1948b, p. 167). Again, why? Similarly there is the "observational versus observed" question. Samuelson is so concerned with "potentially observable phenomenon" – ostensibly a primary motivation for revealed preference theory – and yet he does not seem to be at all interested in the empirical demand analysis. There is also the hanging question of his (early, pre-Houthakker's letter) posit about SARP; he repeatedly said that he had Houthakker's axiom in mind, but could not prove it. Assuming he was being honest, his behavior raises a number of issues. One possibility is that his own mathematical analysis led him to suspect the strong result, even though it was not the non-integrable theory he was looking for. Perhaps it was a case of his mathematical commitments trumping his theoretical commitments, but at this point is just not clear. It is possible that answers to such questions would require more investigation into psychology and personal motivation than is appropriate for historians of economic thought, but they certainly are interesting questions. Moving away from Samuelson per se, and toward the

discussion of the different versions of revealed preference theory discussed in section one: Why has there been no previous historical taxonomy of revealed preference theory? So many different economists working on projects with so little in common and yet all using the same term, as well as so much philosophical and methodological criticism, and yet the historical relationships between the various views has not been explored. Again, there are many questions that need to be investigated. Hopefully this paper has improved our historical understanding of revealed preference theory in general, and Samuelson's relationship to it in particular, but I also hope it is clear that there are still many unanswered questions.

4. Conclusion

The paper has attempted to do two different things. First, it presented a detailed family history of revealed preference theory. I only seriously discussed the main trunk of the family tree – versions that dealt directly with consumer choice theory – but some of the more distant relatives were mentioned in passing. I believe this historical taxonomy is both novel and useful for a wide range of readers: historians of modern economics certainly, but also those interested in methodological issues in contemporary choice theory, as well as economic practitioners in areas where revealed preference theory is applied. The second section is a detailed examination of what Samuelson said over the years (published and unpublished) about revealed preference as well as what others have said about his relationship to it. It is a complex story, but certain consistent themes and tensions emerged among the various writings. Section three examined Samuelson's resistance to the idea that he changed his mind about revealed preference, offered what seems to be the most reasonable rational reconstruction of the texts discussed in the previous sections, and closed with some arguments for why we should not be willing to take any rational reconstruction as the end of discussion.

References

- Afriat, Sidney N. (1967), "The Construction of Utility Functions from Expenditure Data," International Economic Review, 8, 67-77.
- Allen, R. G. D. (1932), "The Foundations of a Mathematical Theory of Exchange," Economica, 36, 197-226.
- Arrow, Kenneth (1959), "Rational Choice Functions and Orderings," Economica, 26, 121-27.
- Arrow, Kenneth J., Block, H. David, and Hurwicz, Leonid (1959), "On the Stability of the Competitive Equilibrium II," Econometrica, 27, 82-109.
- Arrow, Kenneth J. and Hahn, Frank H. (1971), General Competitive Analysis. San Francisco: Holden-Day.
- Bandyopadhyay, Taradas; Dasgupta, Indraneel; and Pattanaik, Prasanta K. (2004), "A General Revealed Preference Theorem for Stochastic Demand," Economic Theory, 23, 589-99.
- Baumol, William J. (1949), "Relaying the Foundations," Economica, 16, 159-68.
- Bernheim, Douglas and Rangel, Antonio (2008), "Choice-Theoretic Foundations for Behavioral Welfare Economics," in The Foundations of Positive and Normative Economics: A Handbook, A. Caplin and A. Schotter (eds.), Oxford: Oxford University Press, 155-92.
- Binmore, Kenneth (2009a), Rational Decisions. Princeton, NJ: Princeton University Press.
- Binmore, Kenneth (2009b), "Interpersonal Comparison of Utility," in The Oxford Handbook of Philosophy of Economics. H. Kincaid and D. Ross (eds.), Oxford: Oxford University Press, 540-59.
- Bronars, Stephen G. (1987), "The Power of Nonparametric Test of Preference Maximization," Econometrica, 55, 693-98.
- Chipman, John S. (1976), "The Paretian Heritage," Cahiers Vilfredo Pareto, 14, 65-171.

Chipman, John S. and Lenfant, Jean-Sébastien (2002), "Slutsky's 1915 Article: How it Came to Be Found and Interpreted," History of Political Economy, 34, 533-97.

Diewert, W. Erwin (1973), "Afriat and Revealed Preference Theory," The Review of Economic Studies, 40, 419-25.

Dorfman, Robert; Samuelson, Paul A.; and Solow, Robert M. (1958), Linear Programming and Economic Analysis. New York: McGraw-Hill.

Evans, Griffith (1930), Mathematical Introduction to Economics. New York: McGraw Hill.

Fostel, Ana; Scarf, Herbert E.; and Todd, Michael J. (2004), "Two New Proofs of Afriat's Theorem," Economic Theory, 24, 211-19.

Gale, David (1960), "A Note on Revealed Preference," Economica, 27, 348-54.

Georgescu-Roegen, Nicholas (1936), "The Pure Theory of Consumer's Behaviour," Quarterly Journal of Economics, 50, 545-593.

Georgescu-Roegen, Nicholas (1954), "Choice, Expectations, and Measurability," Quarterly Journal of Economics, 68, 503-534 [reprinted as Chapter 3 of Georgescu-Roegen's Analytical Economics. Cambridge, MA: Harvard University Press. Page references to reprint].

Gross, John (1995), "Testing Data for Consistency with Revealed Preference," The Review of Economics and Statistics, 77, 701-10.

Gul, Faruk and Pesendorfer, Wolfgang (2008), "The Case for Mindless Economics," in The Foundations of Positive and Normative Economics: A Handbook, A. Caplin and A. Schotter (eds.), Oxford: Oxford University Press, 3-39.

Hands, D. Wade (2006), "Integrability, Rationalizability, and Path-Dependency in the History of Demand Theory," in Agreement on Demand: Consumer Theory in the Twentieth Century, P. Mirowski and D. W. Hands (eds.), Durham, NC: Duke University Press [Annual Supplement to Volume 38 of History of Political Economy], 153-85.

Hands, D. Wade (2009), "Effective Tension in Robbins's Economic Methodology," Economica, 76, 831-44.

Hands, D. Wade (2011a), "Back to the Ordinalist Revolution: Behavioral Economic Concerns in Early Modern Consumer Choice Theory," Metroeconomica, 62, 386-410.

Hands, D. Wade (2011b), "Realism, Commonsensibles, and Economics: The Case of Contemporary Revealed Preference Theory," in Economics from Real: Uskali Mäki and the Place of Truth in Economics. Aki Lehtinen, Jaakko Kuorikoski, and Petri Ylikoski (eds.), London: Routledge, 156-78..

Hands, D. Wade (2011c), "Foundations of Contemporary Revealed Preference Theory" (working paper)

Hausman, Daniel M. (2000), "Revealed Preference, Belief, and Game Theory." Economics and Philosophy, 16, 99-115.

Hausman, Daniel (2008), "Mindless or Mindful Economics: A Methodological Evaluation," in The Foundations of Positive and Normative Economics: A Handbook, A. Caplin and A. Schotter (eds.), Oxford: Oxford University Press, 125-51.

Hicks, John R. and Allen, R. G. D. (1934), "A Reconsideration of the Theory of Value, Parts I and II," Economica, 2, 52-76, 196-219.

Houthakker, Hendrik S. (1950), "Revealed Preference and the Utility Function," Economica, 17, 159-74.

Houthakker, Hendrik S. (1983), "On Consumption Theory," in Paul Samuelson and Modern Economic Theory. E. C. Brown and R. M. Solow (ed.), New York: McGraw Hill, 57-68.

Hurwicz, Leonid (1971), "On the Problem of Integrability of Demand Functions," in Preferences, Utility, and Demand, J. S. Chipman, L. Hurwicz, M. K. Richter, and H. F. Sonnenschein (eds.), New York: Harcourt Brace Jovanovich, 174-213.

Hurwicz, Leonid and Richter, Marcel K. (1979), "An Integrability Condition with Applications to Utility Theory and Thermodynamics," Journal of Mathematical Economics, 6, 7-14.

Hurwicz, Leonid and Uzawa, Hirofumi (1971), "On the Integrability of Demand Functions," in Preferences, Utility, and Demand, J. S. Chipman, L. Hurwicz, M. K. Richter, and H. F. Sonnenschein (eds.), New York: Harcourt Brace Jovanovich, 114-47.

Kihlstrom, Richard; Mas-Colell; Andreu and Sonnenschein, Hugo (1976), "The Demand Theory of the Weak Axiom of Revealed Preference," Econometrica, 44, 971-978.

Little, I. M. D. (1949), "A Reformulation of the Theory of Consumer's Behaviour," Oxford Economic Papers, 1, 90-99.

Metzler, Lloyd A. (1948), Review of Foundations, American Economic Review, 38, 905-10.

Moscatti, Ivan (2007), "Early Experiments in Consumer Demand Theory 1930-1970," History of Political Economy, 39, 359-401.

Moscatti, Ivan and Tubaro, Paola (2011), "Becker Random Behavior and the As-If Defense of Rational Choice Theory in Demand Analysis," The Journal of Economic Methodology, 18, 107-28.

Pareto, Vilfredo (1927), Manual of Political Economy. English translation by Ann S. Schwier, New York: Augustus M. Kelly Publishers, 1971.

Pearce, Kerry A. and Hoover, Kevin D. (1995), "After the Revolution: Paul Samuelson and the Textbook Keynesian Model," in New Perspectives on Keynes, A. F. Cottrell and M. S. Lawlor (eds.), Durham, NC: Duke University Press [Annual Supplement to Volume 27 of History of Political Economy], 183-216.

Pollak, Robert A. (1990), "Distinguished Fellow: Houthakker's Contribution to Economics," Journal of Economic Perspectives, 4, 141-56.

Richter, Marcel K. (1966), "Revealed Preference Theory," Econometrica, 34, 635-45.

Richter, Marcel K. (1987), "Revealed Preference Theory," in The New Palgrave: A Dictionary of Economics, Vol. IV, J. Eatwell, M. Milgate, and P. Newman (eds.), London: Macmillan, 166-70.

Rosenberg, Alexander (1992), Economics – Mathematical Politics or Science of Diminishing Returns? Chicago: University of Chicago Press.

Samuelson, Paul A. Papers. Rare Book, Manuscript, and Special Collections Library, Duke University, Durham, North Carolina.

Samuelson, Paul A. (1938a), "A Note on the Pure Theory of Consumer's Behaviour," Economica, 5, 61-71.

Samuelson, Paul A. (1938b), "A Note on the Pure Theory of Consumer's Behaviour: An Addendum," Economica, 5, 353-54.

Samuelson, Paul A. (1938c), "The Empirical Implications of Utility Analysis," Econometrica, 6, 344-56.

Samuelson, Paul A. (1938d), "The Numerical Representation of Ordered Classifications and the Concept of Utility," Review of Economic Studies, 6, 65-70.

Samuelson, Paul A. (1947), Foundations of Economic Analysis. Cambridge, MA: Harvard University Press.

Samuelson, Paul A. (1948a), "Consumption Theory in Terms of Revealed Preference," Economica, 15, 243-53.

Samuelson, Paul A. (1948b), "International Trade and the Equalisation of Factor Prices," Economic Journal, 58, 165-84.

Samuelson, Paul A. (1950), "The Problem of Integrability in Utility Theory," Economica, 17, 355-85.

Samuelson, Paul A. (1955), "Linear Programming and Economic Theory," in The Collected Scientific Papers of Paul A. Samuelson, Vol. I, J. E. Stiglitz (ed.), Cambridge, MA: MIT Press, 1966, 493-504.

Samuelson, Paul A. (1963), "Discussion," American Economic Review, 53, 231-36.

Samuelson, Paul A. (1964), "Theory and Realism: A Reply," American Economic Review, 54, 736-39.

Samuelson, Paul A. (1965), "Professor Samuelson on Theory and Realism: Reply," American Economic Review, 55, 1164-1172.

Samuelson, Paul A. (1972), "Maximum Principles in Analytical Economics," American Economic Review, 62, 249-62.

Samuelson, Paul A. (1983), Foundations of Economic Analysis, Enlarged Edition, Cambridge, MA: Harvard University Press.

Samuelson, Paul A. (1998), "How Foundations Came to Be," Journal of Economic Literature, 36, 1375-86.

Sen, Amartya K. (1973), "Behaviour and the Concept of Preference," Economica, 40, 241-59.

Sen, Amartya K. (1993), "Internal Consistency of Choice," Econometrica, 61, 495-521.

Sen, Amartya K. (1997), "Maximization and the Art of Choice," Econometrica, 65, 745-779.

Slutsky, Eugene E. (1915), "Sulla Teoria del Bilancio del Consonatore," Giornale degli Economisti, 51, 1-26. Translated in Readings in Price Theory, G. J. Stigler and K. E. Boulding (eds.), Homewood, IL: Richard D. Irwin, 1952, 27-56.

Stigler, George J. (1948), Review of Foundations, Journal of the American Statistical Association, 43, 603-05.

Varian, Hal (1982), "Nonparametric Test of Models of Consumer Behavior," Review of Economic Studies, 50, 99-110.

Varian, Hal (1985), "Non-Parametric Analysis of Optimizing Behavior with Measurement of Error," Journal of Econometrics, 30, 445-58.

Varian, Hal (2006), "Revealed Preference" In M. Szenberg, L. Ramrattan, and A. A. Gottesman (eds.), Samuelsonian Economics and the Twenty-First Century, Oxford: Oxford University Press, 99-115.

Wald, Abraham (1951), "On Some Systems of Equations of Equations of Mathematical Economics," Econometrica, 19, 368-403 [translation of 1936 article appearing in Zeitschrift für Nationalökonomie].

Weintraub, E. Roy (1991), Stabilizing Dynamics: Constructing Economic Knowledge. Cambridge: Cambridge University Press.

Wong, Stanley (2006), The Foundations of Paul Samuelson's Revealed Preference Theory. London: Routledge [originally published in 1978].