HOUSEHOLD CONSUMPTION: RESEARCH QUESTIONS, MEASUREMENT ISSUES, AND DATA COLLECTION STRATEGIES [1]

Luigi Pistaferri
(Stanford University, NBER, CEPR, and SIEPR)

Department of Economics
Stanford University
Stanford, CA 94305
Tel. +1 (650) 724-4904
Fax: +1 (650) 725-5702
E-mail: pista@stanford.edu

ABSTRACT

In this essay I summarize what I see as the most important academic and policy issues related to the study of consumption behavior in the US (and elsewhere). I discuss the type of data researchers working on these topics have available, the problems they encounter, and how the "big data revolution" is rapidly changing access and availability of data on household spending. Finally, I discuss the arguments in favor and against starting a new national panel survey collecting consumption-related information as opposed to enhancing or improving the existing surveys.

Keywords: Consumption, Survey Data.

[1] Manuscript prepared for the "Workshop on New Household Panel Survey". Thanks to an anonymous referee, Tullio Jappelli, Bob Schoeni, the discussant David Johnson and the members of the Steering Committee for useful comments.

1. INTRODUCTION

Household consumption is the most important component of GDP. In the US, for example, it accounts for about 70% of GDP. Consumption shares of GDP in countries at a similar level of development are slightly smaller, but still quite high (60% in Germany, 66% in the UK, 58% in Canada - to cite just a few). Consumer confidence (as reflected in purchase decisions) is one of the most crucial indicators for judging the health of an economy, and swings to the consumer confidence index are key to predict the speed of a recovery or the depth of a recession. Knowing what determines consumption decisions, and how consumers change their behavior over the life course, over the business cycle, and in response to changes in their income, asset holdings, and personal circumstances (such as unemployment, retirement, disability, demographic shocks, etc.) are important research questions, and hence it becomes vital to have appropriate data, at the micro and macro level, for being able to answer such questions in a statistically reliable and precise way.

Not surprisingly given its quantitative relevance, studying the consumption decisions and the spending habits of American households has attracted the attention of a large and active research community, both from a macroeconomic perspective (e.g., to predict the response of consumption to tax stimulus packages, etc.) as well as from a microeconomic perspective (e.g., to predict how consumers respond to changes in pricing strategies, marketing and advertising campaigns, to investigate brand loyalty, etc.).

In this essay I will first try to summarize what I see as the most important academic and policy issues related to the study of consumption behavior in the US (and elsewhere). I will then discuss what type of data researchers working on these topics have available, the problems they encounter, and how the "big data" revolution is rapidly changing consumption data access and

availability, and how we may expect to change it even further in the future. Finally, I will discuss the case in favor and against starting a new national panel survey collecting consumption-related information as an alternative to enhancing or improving the existing surveys containing information on household spending.

2. RESEARCH ON CONSUMPTION

Traditionally, consumption decisions have been studied by macroeconomists and IO/marketing researchers. To keep this essay within reasonable limits, I will focus primarily on work of the first type.

Macroeconomists are usually interested in understanding how consumers respond to tax or welfare policy reforms that change the level of economic resources or household expectations about the path of future resources. The theory that modern economists use to predict such responses (the life cycle permanent income hypothesis of Modigliani-Brumberg and Friedman) suggests that the response will depend on the nature (anticipated vs. unanticipated) and persistence (temporary vs. long-term) of the changes in economic resources. Enriched to include salience or inattention, adjustment costs, or various other market frictions, also the size, the sign, the medium, and the context in which such changes occur may play a role.

Fig. 1 (extending Jappelli and Pistaferri, 2010) provides a simple taxonomy. The theory predicts that rational, risk-averse consumers use all the information they have available to form optimal consumption plans constrained only by their lifetime resources (and not by their current resources, as credit markets are assumed to allow consumers to frictionlessly move resources from periods of abundance to periods of scarcity). Consumption plans are revised only when new

information about the future arrives. Hence, we should expect only unanticipated changes in resources ("shocks") to induce deviations from optimal consumption plans. Moreover, temporary shocks should generate a smaller consumption response than a more persistent shock, as it is easier to borrow against the former than against the latter. In life-cycle models, this response heterogeneity is reduced as the consumer's horizon shrinks (in the last period of life, there is no meaningful distinction between a transitory and a permanent shock).

[Fig. 1 here]

Adding frictions makes the consumer's response to changes in economic resources less clear-cut and may induce asymmetric behavior. For example, in models with liquidity constraints, positive transitory shocks induce strong consumption responses (as they relieve consumers of the borrowing constraints they face), while negative changes do not. Similarly, consumers may respond to anticipated income increases (but not to anticipated declines), as credit markets imperfections may prevent individuals to borrow against positive income growth (but nothing prevents consumers from saving).

Adjustment costs may also change theoretical predictions, irrespective of liquidity constraints. For example, if the cost of adjusting consumption upwards in anticipation of a future income increase is high, consumers may not respond to small anticipated changes in income, while they may do it if the anticipated income change is large. Kaplan and Violante (2013) explain some of the excess response to tax rebates found by e.g., Johnson, Parker and Souleles (2006), with a model in which there are "wealthy hand-to-mouth" consumers, i.e., consumers who are notionally wealthy (have housing or pension wealth, for example) but in practice act as constrained consumers because most of their wealth is illiquid or can only be liquidated at high cost. This highlights the importance of observing in the data not only consumption and income

information but also the level and composition of one's net worth if the goal is to predict the response of consumers to tax rebates, for instance.

As Fig. 1 suggests, consumer responses may also depend on the *context* in which the change in income occurs (i.e., the household's leverage position, the stage of the life cycle or the business cycle, etc.), or - if inattention costs are important - it may also differ depending on the *medium* with which it is implemented (e.g., a rebate check in the mail may be more salient than a reduction in the payroll tax).[2]

Most applied research has been focused on estimating the response of consumption to the various changes outlined in Fig. 1. In this line of work, researchers are primarily interested in measuring marginal propensities to consume or consumption elasticities. These parameters are important both to understand the validity of existing economic theories of consumption behavior as well as for policymakers weighting the countercyclical benefits of stimulus packages (which depend on the size of the multiplier) against the budgetary costs of the policy. The empirical challenge is to find exogenous and credible measures of anticipated vs. unanticipated resources, transitory vs. permanent shocks, etc. Some researchers use demographic shocks (family formation and dissolution, arrival or departure of kids, and so forth), or watershed events such as disability, unemployment, or retirement, to identify such exogenous events. Jappelli and Pistaferri (2010) survey this vast literature. In general, "excess" response of consumption to certain changes in resources or certain events is informative about the extent of insurance that consumers have available, and hence have direct welfare implications.

[2] It would be interesting to compare the effectiveness of the 2011-12 Obama payroll tax cut with traditional tax rebate interventions implemented in the past. Sahm, Shapiro, and Slemrod (2012) compare the 2009 change in tax withholding resulting from the enactment of the Making Work Pay Tax Credit Act with the 2001 and 2008 "one-shot" tax stimulus payments and find that the 2009 reduction in withholding increased consumer spending by half of what achieved with the 2008 one-shot payment (13% vs. 25%).

Recently, most attention has been paid to the issue of response heterogeneity, which calls for greater "targeting" of policies as opposed to generalized interventions if the goal is maximization of the policy's impact. To address MPC heterogeneity, Shapiro and Slemrod (1995, 2003) and Jappelli and Pistaferri (2014) use self-reported marginal propensities to consume in response to actual (or, in the latter case, hypothetical) income changes, while Misra and Surico (2012) use a sophisticated econometric methodology. It should be noted that in most cases the estimated MPCs are not necessarily structural parameters in the sense of Lucas (1976). Recent work by Martin Browning and co-authors (2010) could be used to estimate the underlying "structural" sources of heterogeneity in the MPC (whether coming from preferences, technology, or constraints), which better lends itself to policy analysis.

While most of the literature has looked at changes in income, consumers' resources ("cash on hand") also include wealth, and household wealth changes over time for exogenous reasons (changes in asset prices) as well as for endogenous reasons (portfolio choices). The period surrounding the recent financial crisis, with its large increase in housing and stock market wealth, followed by the bursting of the housing bubble and the consequent destruction of wealth, has fed enormous interest on the importance of the so-called "wealth effect" (i.e., how consumption responds to exogenous changes in net worth). Related to this is the relevant policy question of what effects monetary policy may have on consumption levels (and consumption inequality, see Coibion et al., 2012). Monetary policy may affect consumption in a variety of ways, but chiefly through two channels: (a) by changing real interest rates (and hence inducing intertemporal shifts between consumption and savings), and (b) by relieving/amplifying the effect of borrowing constraints by the way of "easy money" policies or credit market crunches (see Mian, Rao and Sufi, 2013). For example, most borrowing constrained households may be

able to better smooth their consumption by refinancing their mortgages, either through equity extractions or enjoying the flow effect associated with lower mortgage interest payments. How important these channels are is still not clear. For example, there is not much evidence on how the quantitative easing policy enacted by the FED has affected consumption behavior. It is also notoriously difficult to estimate the key elasticity of intertemporal substitution parameters, as we do not have much heterogeneity in interest rates across consumers and these variables typically do not vary much over time.

Most studies have tried to infer the importance of borrowing constraints by estimating Euler equations. The problem with using this framework is that the Euler equation may not necessarily be violated even in the presence of borrowing constraints, and that it is difficult to disentangle liquidity constraints from precautionary savings (both predict a growing consumption profile over time). Given the mixed evidence for liquidity constraints in tests that use the Euler equations, an alternative avenue is to search for the effect of credit market imperfections on durable consumption and housing, such as in Gerardi et al. (2010), Gross and Souleles (2002), Adams, Einav and Levin (2009), or Attanasio, Goldberg and Kyriazidou (2008). Another route is to use direct indicators of liquidity constraints, such as those used by Jappelli (1990).

In models with uncertainty and prudent consumers, changes in perceived uncertainty (i.e., second and higher moments of the individual distribution of future resources) are as important as changes in expectations (first moments). There is a growing literature on the importance of "uncertainty shocks" (see Bloom, 2014) for explaining consumer confidence or the fact that consumer may refrain from purchasing goods (especially durables) if uncertainty is high (reflecting precautionary behavior or the combined effect of caution and adjustment costs). This can slow down recoveries and affect the impact of public policies. From a data-related point of

view, one key difficulty is that it is hard to identify observable and exogenous source of risk that varies significantly across the population (Browning and Lusardi, 1996).

Related to the study of individual second moments is the study of shifts in the distribution of consumption both over time and across groups. Interest in the study of inequality (or mobility, its dynamic counterpart) is indeed soaring, as reflected in academic work, policy debates, and articles in the popular press (and even movies![3]). But while it is well known that wage and earnings inequality have increased dramatically in the US over the last 35 years, research on consumption inequality, possibly due to measurement issues (discussed below), is less conclusive. Yet it is key, as from a policy perspective it is important to assess whether the well-documented increase in wage inequality has led to an increase in disparities in welfare, and consumption is believed to be a better measure of welfare than income.[4] Early research pointed to consumption inequality rising much less than wage inequality, or even remaining stable (Krueger and Perri, 2006); more recent research shows a much closer tracking between inequality in wages/income and inequality in spending (Attanasio and Pistaferri, 2014; Aguiar and Bils, 2011; Attanasio, Hurst and Pistaferri, 2013).[5] An important topic for future work is to understand the sources of these discrepancies.

[3] See, e.g., Robert Reich's documentary "Inequality for all".

[4] In the models discussed earlier, consumption depends on long-run (or permanent) income, not transitory income; taxes and transfers and saving and borrowing can drive a wedge between current and permanent income. Meyer and Sullivan (2009) discuss why a consumption-based measure of poverty has the same advantages but fewer disadvantages than existing measures of poverty that are based on income. Clearly, for consumption to be a more reliable measure of welfare than income we need to have access to good and reliable measurement of household consumption. These issues are discussed in the next section.

[5] This revision parallels the one involving the extent that income inequality has been transitory or permanent. Earlier work by Gottschalk and Moffitt (1994), based on survey data, concluded that transitory volatility had been substantial, while more recent work using administrative data, such as De Becker et al. (2103) or Guvenen et al. (2014) finds that most of the increase in wage and earnings variance has been structural, or of more permanent nature.

The study of consumption inequality can be used to test theoretical predictions by looking at its evolution over the business cycle and over the life cycle. Fig. 2 (using bi-annual data from the PSID) shows that consumption inequality increases before the Great Recession (consumption growth at the top is faster than at the bottom), but then it shrinks during the economic bust (consumption growth at the bottom is now faster, or to be more precise less slow, than at the top). This finding may appear puzzling at first – indeed, wealthy individuals are expected to have better tools to smooth their consumption during rough periods. However in the context of the Great Recession, another effect might dominate the consumption behavior of wealthy individuals. As shown in Wolff, Owens and Burak (2011), the recession was characterized by large wealth destruction suffered by rich individuals. This suggests that they lost a large fraction of their "buffer" wealth, which was supposed to be used exactly for the purpose of smoothing their consumption during the "rainy days" of the recession. To restore their "buffer stock" these individuals may have needed to save more, which would have slowed down their consumption growth.

[Fig. 2 here]

Another issue of importance is how to explain cohort shifts in lifetime consumption inequality, as shown in Fig. 3. Here we are using cohort-specific consumption inequality data obtained using the Attanasio and Pistaferri (2014) "backward imputation" strategy. The most striking aspect of this graph is that not only consumption inequality for more recent cohorts is higher at any point of the life cycle relative to previous generations at the same age (including, crucially, at the point of entry in the labor market), but also that the growth in consumption disparities seems to have grown faster within a cohort.

[Fig. 3 here]

Finally, intergenerational and intra-person mobility in consumption is understudied, primarily due to lack of appropriate data. Once more, this analysis may be more informative than the study of mobility in incomes given the interpretation of consumption as a better measure of welfare. The study of consumption mobility across generations is informative about whether there is persistence in the welfare rankings across generations; interpersonal consumption mobility informs us about persistence of welfare rankings within generations. In Fig. 4 we present some simple estimates of interpersonal mobility using the quarterly rotating panels of the CEX. First, we construct an empirical transition matrix of consumption and measure mobility using a simple Shorrocks' index.[6] This is an approximate measures of the fraction of individuals moving across the distribution; a higher value of the index (which ranges from 0 to 1) is associated to a higher degree of mobility from one year to the next. In Fig. 4 we plot the monthly Shorrocks' indexes as well as the local regression smoothed version (the solid line). Overall, there seems to be a general declining trend in the amount of mobility in consumption. What can explain this trend? An improvement in the amount and nature of insurance available to households (such as through a process of financial market liberalization that makes credit more easily available and more widespread as it happened during the 1990s and part of the 2000s) may have contributed to a reduction in the amount of mobility by attenuating the response of consumption to shocks. The idea that credit market development in the US has helped households in smoothing shocks more efficiently is also advanced by Krueger and Perri (2006). The fact that mobility rises (or at least stops declining) during the Great Recession is consistent with the worsening of opportunities for smoothing shocks offered by credit markets.

[6] The index is defined as $S=(q-\Sigma_i p_{ii})/(q-1)$, where q is the number of quantiles in which the distribution has been divided, and p_{ii} is the proportion of individuals who remain in quantile i between two adjacent periods. In our analysis, we work with deciles (q=10).

[Fig. 4 here]

In an important series of paper, Aguiar and Hurst (2007) have highlighted the importance of home production for explaining some consumption "puzzles" (notably, the fall of consumption at retirement). With home production and efficient use of time, some consumption is produced in-home rather than purchased on the market; moreover, time can be used to search for lower prices. The most important consequence of this work is that the observed life-cycle time allocation implies a consumption series that differs markedly from the expenditure series. The heterogeneity in the price faced by consumers for seemingly homogenous goods has also important implications, for say, consumption inequality measurement. "Real" consumption inequality should account for the fact that low-income (or low-spending) households may face different prices for the same goods (i.e., because their value of time is different, or because they live in low cost-of-living areas, etc.). Similarly, when attention is limited to food, it could be important to distinguish between spending and consumption by focusing on the nutrition value of the food consumed, as well as the health content thereof. Research on this topics are limited by the fact that time-use data are rarely collected alongside spending data.

Recently, some authors have looked at the implications of home production on the supply side. Kaplan and Menzio (2013) consider a model in which firms have market power that depends on how hard people shop for deals; unemployed individuals are more likely to have the time to do so. If a firm hires a worker, this generates positive externalities on other firms, both because of a demand effect (employed individuals consume more) and a supply effect (employed individuals have less time to shop for deals). Similar interesting "supply" effects have been observed by Hastings and Washington (2010) (grocery shops serving the population of welfare recipients may change strategically their pricing in order to reflect the timing of welfare

receipt),[7] and by Della Vigna and Pollet (2007), who noticed that cohort size fluctuations (i.e., the baby boom cohort, etc.) generate predictable shifts in demand for certain goods (toys, life insurance, nursing homes, etc.), and firms can adapt strategically to such forecastable changes. Finally, Hendal and Nevo (2013) show that firms appear to use sophisticated forms of intertemporal price discrimination, i.e., they use different pricing strategies for consumers who stockpile (i.e., Costco shoppers) and those who do not.

Another lively branch of consumption research is intrafamily allocation of consumer goods. Models that assume the household behaves unitarily are replaced by collective models of behavior, in which individuals in the household have distinct preferences and among whom a collective decision process takes place. These models can be used to understand why families are formed or dissolved, female labor supply or human capital investment driven by "strategic" considerations, the composition of household consumption, etc. From a data point of view, the most complicated aspect is assignment of private and public consumption (in order to compute individual welfare, etc.). Data sets rarely have a level of detail that allow such assignments, and indeed it is not inconceivable that research on these topics has lagged behind due to lack of appropriate data. Recently, Cherchye, De Rock and Vermeulen (2012) have collected a brand new data set for the Netherlands that allows them to identify and separate private good from public good consumption for all household members, and moreover it contains detailed time use information. There are no such data available for the US.

Finally, there is an enormous and growing literature by "behavioral" economists documenting empirical failures of the traditional model (see Della Vigna, 2009, for a survey). For example, consumers appear to exhibit inertial behavior, as exemplified by the literature on

[7] This assumes that grocery stores recognize the strength of borrowing constraints and hence the inability of welfare recipients to smooth consumption from welfare check to welfare check.

the power of default clauses. In these models, individuals appear to "stick" to default clauses even in the absence of significant cost of switching and in the presence of substantial benefits from switching. Consumers also appear to exhibit hyperbolic discount rates, as evidenced by procrastination of "good behavior" (saving for retirement or in response to expected increases in the interest rate being chief examples). Individuals may be sensitive to how information is "framed" or presented – there is extensive lab experiment or real-life evidence of loss aversion or of the importance of salience regarding prices, for example. Moreover, mental accounting may explain the (seemingly puzzling) differential response to certain types of income changes (such as one-shot tax rebates vs. fundamental tax reforms). There is the usual problem of disentangling how much of the empirical failure of the traditional model is due to "behavioral" considerations and how much to the operation of "frictions". Data that would permit us identification of these two phenomena would be extremely valuable. In general, we have little information on the extent of adjustment costs faced by individuals.

There is also a growing literature on the effect of interdependent preferences, or network/peer effects, on behavior. The study of network effects has been primarily concentrated in fields like education (test score effects), public finance (saving for retirement choices, welfare use), and labor economics (productivity). While traditional consumption theory assumes that consumers behave atomistically, there have been many historical attempts to introduce interdependent preferences into traditional models (the relative consumption hypothesis *à la* Duesenberry, the conspicuous consumption hypothesis of Veblen, as well as the keeping-up-with-the-Joneses hypothesis). This should be a fruitful area of research in the future, as it may also have important policy implications (to study distortions in intratemporal or intertemporal consumption choices,

as well as for the possibility of social multiplier effects).[8] De Giorgi, Frederikssen and Pistaferri (2014) study consumption network effects with administrative tax records data for Denmark. Bertrand and Morse (2014) use CEX data to test whether in the last decades the consumption of poorer households has been driven by consumption behavior of households at the top of the wealth or income distribution. This idea is behind Rajan's narrative (2010), claiming that the financial crisis was partly fueled by over-borrowing at the bottom to keep up with the (rich and getting richer) Joneses.

3. EXISTING CONSUMPTION AND SPENDING SURVEY DATA

In the US, researchers have available several surveys containing information on household expenditure. These surveys are useful to paint a picture of the spending habits of Americans, but as I will discuss next, are not without problems.

3.1.A (brief) description of available survey data

The Consumer Expenditure Survey (CEX) is the only data set with comprehensive and detailed information on household expenditure and its various components.[9] Available on a continuous basis since 1980, it is used by the Bureau of Labor Statistics (BLS) to form weights

[8] Hence estimation of MPCs may also depend on the extent of network relations.

[9] The Personal Consumer Expenditure data, collected by the Bureau of Economic Analysis, is the other source of data on household spending (and forms the basis for the national accounts, or NIPA data). For most categories, PCE spending is estimated using a "commodity-flow" method. In this method, one starts with the value of domestic output based on data from the census of manufacturers (i.e., looks at the value of manufacturers' shipments, etc.). Next, domestic consumption (denominated in producers' prices) is estimated by adding imports and subtracting exports and changes in inventory. Finally, the value of consumer purchases is converted from producers' prices to purchasers' prices by adding wholesale margins and taxes, transportation costs, and retail margins and taxes.

that go in the computation of the Consumer Price Index (CPI) (and for other minor matters as well).[10] The CEX is composed of two distinct surveys, the Interview survey (where spending information is by three-month recall), and the Diary survey (where spending is collected by filling a 2-week diary). Respondents in the Interview survey are sampled every 3 months (for a total of 5 times, although data for the first interview are not released because they are merely preparatory), while those in the Diary survey are sampled only once. The two surveys cover different consumption items, with some overlap.

The other data set that is widely used by academic researchers to study consumption behavior is the Panel Study of Income Dynamics (PSID). Available on an annual basis from 1968 to 1997, and on a bi-annual basis after that, the PSID's initial goal was to permit social scientists to study income dynamics (and poverty) between and across generations. For this reasons, information on consumption was ancillary. Until 1997, the PSID collected information only on a few consumption items – food (at home and away from home), rent paid by renters, and (very occasionally) utility payments. Starting with the 1999 wave, however, the PSID has began collection of information on a larger range of items (covering about 70%-90% of the items collected in the CEX). Respondents typically report spending for broad categories, with the reference period being (with some exceptions) the previous calendar year.

Researchers have also used other data sets to study consumption, but they all come with some limitations. For example, the Health and Retirement Survey-CAMS module and the American Life Panel try to collect comprehensive information on consumption with a few targeted questions. The Survey of Income and Program Participation (SIPP) collects some limited consumption information in a very ancillary manner. The Residential Energy

[10] The CEX was also collected, with different sampling strategies, in 1960-61 and 1972-73.

Consumption Survey, as the name suggests, is primarily about utilities and ownership of some durables. Similarly, the American Community Survey reports only information on housing expenses (rent, utilities, home insurance, mortgage payments).

Finally, the "big data" revolution of the last decade has brought several data sets containing information on spending to the fore. I will discuss the challenges and opportunities offered by these new data sources in one of the sections below.

3.2.Problems with existing survey data on consumption

While the CEX is in principle the data set of reference for researchers interested in studying consumption behavior, and it has indeed been used to study some of the topics discussed in the previous Section, there are several problems with the survey. The most relevant one is that there is evidence of growing detachment of average consumption as measured in the CEX from NIPA (National Accounts) aggregates. In particular, there is serious under-reporting. Fig. 5, drawn from Heathcote, Perri and Violante (2010), shows that there is growing detachment for non-durable consumption (in particular, food consumption – the sum of food at home, food away from home, and alcohol beverages), while the detachment is less serious for durable consumption (although the coverage ratio declines over time) and services.[11]

[Fig. 5 here]

[11] The results are less dramatic when an effort is made to compare items that are measured "similarly" in the two data sets. Passero, Garnero and McCully (2014) make such comparison and conclude that "non-durables are most alike for the CE and PCE with about 93 percent of total non-durable expenditures identified as comparable within the CE and within the PCE." However, their conclusion is also that "focusing on comparable goods and services only, CE to PCE ratios have steadily decreased. For total comparable goods and services, CE to PCE ratios decreased from 84 percent for 1992 to 74 percent for 2010. The greatest decline in CE to PCE ratios is for durables, with a decrease of 24 percentage points. Ratios for comparable services dropped the least, with a percentage decrease of 10 percentage points".

Aguiar and Bils (2012) report that the hypothesis of classical measurement error in CEX consumption data appears inconsistent with saving data (by income levels) reported elsewhere in the survey. Other papers also show that that there is more severe understatement of spending among rich households and that survey participation among high income households has declined over time and that there is greater under-reporting for luxuries (such as jewelry) that the rich consume to a larger extent.

Bee, Meyer and Sullivan (2012) assess the performance of the CEX surveys on a good-by-good basis and report three notable findings. First, in general the Interview survey performs better than the Diary survey in matching NIPA aggregates for some categories. Second, the coverage ratios are excellent for some goods (food at home) and have not changed appreciably over time; on the other hand, the coverage ratios for other items (such as clothing) are low and declining. Finally, durable stocks (cars) and durable purchases appear to be reported sufficiently well. In terms of general representativeness of the survey, they conclude that "based on observable characteristics, the [CEX] appears to be fairly representative, although there is strong evidence of under-representation at the top of the income distribution and under-reporting of income and expenditures at the top." Another worrying trend is the increase in non-response.

Blundell, Pistaferri and Saporta-Eksten (2012) show the performance of the redesigned PSID consumption measure against NIPA data, as reported in Fig. 6, where we plot 2-year growth rates in the PSID and in NIPA. As is visible from the graph, the matching is sufficiently good (especially considering that collection of consumption data is not a priority of the data set, that the two definitions of consumption are not strictly comparable, and that the timing of the data is not perfectly synchronized), but far from perfect.

[Fig. 6 here]

The problems associated with the CEX have intensified the call for a redesign of the survey, which is currently underway following the recommendations of a specially appointed NAS panel (the "Gemini" project). The PSID new consumption data have proved able to capture some broad aggregate trends, despite their lower scale and scope, but come from a small sample, where representativeness can be a problem, and where coverage of the goods that households purchase is less than 100%.

The discussion in Section 2 of the topics of interest in the consumption literature suggests that while the CEX and the PSID are valuable sources of information, they are far from ideal.[12] For example, the CEX panel is short. Longitudinal information is important to understand dynamic consumption behavior and seems a *must* for economists working with life cycle models.[13] Moreover, the frequency of consumption and income (and wealth) data is not synchronized, making hard to study questions related to consumption response to changes in resources (unless additional "one-time-only" information is made available, such as that used by Johnson, Parker and Souleles, 2006). Finally, with the exception of cars, there are no information on durables stocks – and information on durable purchases must confront the incidence of zeroes coming from infrequent purchases (which is even more severe in the Diary Survey, for obvious reasons).[14] The scope of the CEX is also limited in other dimensions. Geographical information is sketchy (and censored for people living in small states); there is no information on networks that households may belong to; the distinction between individual and public consumption is

[12] Of course, most questions have been approached using the existing data, but with mixed results and large noise.

[13] There is a large empirical literature obviating the absence of panel data by working with repeated cross-sections, as in the seminal work of Browning, Deaton and Irish (1985). While these data follow cohort of individuals, they miss the truly idiosyncratic heterogeneity within a given cohort. Since there is growing evidence that the rise in income inequality has been as much within as between groups, the inability of cohort data of analyzing such form of heterogeneity is not a negligible issue.

[14] Reconstructing the stock of cars is not easy, as consumers report the year, type, and make of the cars they own – not their resale values, which must instead be estimated from specialized publications (such as the Kelley Blue Book, see Padula 1999).

very limited (only for clothing there is a gender/adult vs. kids categorization); there is no information on time use; and there is no possibility to link CEX households with IRS, Social Security records, or other administrative data. Finally, the CEX does not ask any subjective questions to respondents that may be used to measure preferences, constraints, or expectations. These questions can be important to test theoretical predictions. For example, Pistaferri (2001) shows that the availability of a panel of subjective expectations of future income and actual income realizations allows *point identification* of income shocks, permitting a direct test of the proposition that consumption should respond more to permanent than transitory income shocks. Subjective expectations can also be used to measure risk, an important variable that is extremely difficult to quantify in data based on realizations.[15]

The PSID has a longer longitudinal span and its consumption data, while limited, have proved able to capture some broad aggregate trends, as we documented above (see Blundell, Pistaferri and Saporta-Eksten, 2012). But more research is needed on the ability of the redesigned PSID to capture macro trends.[16] The PSID sample is small and it does not cover all consumption goods that households consume. Moreover, while in the CEX the most important service from durables components are directly elicited (services from housing or imputed rent) or can be constructed from available data (services from cars), both are absent in the PSID.[17]

[15] Against this background, there is the usual problem of how much we should trust "what people say" as opposed to "what people do" – see Manski (2004) for a general discussion of the usefulness of subjective expectations data and of expectation data collection strategies.

[16] For example, Blundell, Pistaferri and Saporta-Eksten (2012) notice that many questions have item non-response (missing data) and it is not clear whether missing records reflect truly no spending or refusals.

[17] In principle, imputed rent can be reconstructed on the basis of the self-reported value of one's housing, but this may introduce various concerns about measurement errors. In the survey people also report the number of cars they own. Since 1999, the PSID has also asked respondents to report the make, model, and year of up to three vehicles. The vehicle make and year are on the public use file. The model is only available via restricted use contracts because it can jeopardize confidentiality. Hence, with some effort it is possible to estimate the imputed services from the owned vehicles.

Finally, the PSID shares with the CEX the limitation of lacking information on social networks, subjective expectations, time use, and within-household consumption composition.

4. IS A NEW NATIONAL HOUSEHOLD SURVEY NEEDED?

The problems with measuring consumption in traditional surveys and the paucity of information needed to address convincingly most of the questions of interest to economists or other social scientists discussed in Section 2 have spurred interested in starting a brand new survey. A key question is whether we need a new survey collecting information on what Americans consume or rather a redesign of the existing surveys.

I discuss three possible routes in what follows. First, possible improvements or enhancements of existing surveys (the CEX and the PSID). Second, starting a new survey from scratch. Third, introduction of novel data collection strategies (which may be interspersed with the two previous routes, i.e., one could start a new survey *with* novel data collection strategies). I discuss pros and cons of these three routes. I stress that my opinions here reflect the point of view of an academic end-user of data, so it comes with the obvious caveats of not considering all the benefits and costs aspects of the three proposals, especially on the "practical implementation" phase.

4.1.Improving the existing surveys

To improve existing survey data we need to have a good understanding of what works and what doesn't, and we need to have a good sense of priorities. Thus, we need at the same time to

improve measurement of household spending and to increase the scope ("information coverage") of the surveys in order to be able to use the data to answer questions of interest to economists and policymakers.

Consumption data in complex economies are difficult to collect and there is evidence of substantial mis-measurement. As I tried to discuss above, the evidence is more nuanced than it may appear at first, as consumption is made of the sum of various components which may be collected with different strategies, and hence each components may have different measurement problems.

What lessons have we learned from decades of comparing aggregate statistics from the CEX (or PSID) to NIPA data? Which goods are better measured? Are retrospective questions better suited than diary collection methods? We should heed the lessons we have learned and use the existing evidence to improve measurement. For example, Bee, Meyer and Sullivan (2014) show that in the Interview survey of the CEX certain goods have extremely high coverage ratios and that such ratios have not changed appreciably over time.[18] Stable precision of data implies that we should continue to get information on those goods using similar methodologies. But for other goods, where coverage ratios are small or falling, we need some understanding of where failure lies and some experimenting of how it can be overcome. Note that most of what we say for existing surveys applies equally well to brand-new surveys – the way we collect data in new surveys should take into account the lessons learned from measuring goods in existing surveys. Slightly differently, the idea of changing data collection strategies within an existing survey (i.e.,

[18] One way of improving coverage ratios and reduce measurement error in the Diary section is to provide respondents with scanners or internet tablets, as argued by the NAS panel on the redesign of the CEX. However, Diary data collection methods are becoming increasingly less reliable due to secular changes in the way people purchase goods: consumers now more frequently purchase goods in bulk, over the internet, make fewer trips to grocery stores per time unit, or "consume" goods without actually paying for it (P2P entertainment streaming). It is possible to devise econometric methodologies to account for increasing infrequency of purchase.

asking a global "spending on clothing" question, instead of multiple questions on "women shirts", "men shirts", "children shirts", etc.) is unlikely to be helpful, as it creates serious comparability problems with aggregate data from previous waves.

Above, I mentioned two other problems with the existing data sets. The first is information coverage. The CEX, understandably, contains a wealth of information about spending, some limited information on income and wealth, and relatively poor information on other labor market or socio-demographic aspects. The PSID has its own coverage limits. A lot of the economic questions that I discussed in Section 2 can be answered only imperfectly or not at all with the existing amount of information. It would be desirable to add new information to the CEX and the PSID, but how? The problem is that the CEX is already quite long and asking survey respondents to provide additional information may result in an increase in item non-response due to survey fatigue. As for the PSID, different sections of the survey are of different interest to social scientists and assessing priorities is important – which sections could be sacrificed to make room for new sets of questions? One possible way to add information at a reduced cost is by introducing "special modules", as done for example in the SIPP. But unlike the SIPP, which is conducted every 4 months, the PSID goes in the field every two years. So modules should try to be compact, rotating, and ask questions to randomly selected subsamples. For example, half of the sample could be asked subjective expectation questions; the other half, subjective questions on preferences. Perhaps more experimenting is possible with the CEX given its 3-month rotating structure.

The second issue with the existing surveys is their small sample sizes. A large sample size increases precision of estimates and allows researchers to reduce type I and type II errors when conducting inference. A large initial sample also allow researchers to be less conservative when

selecting their final estimation sample. Finally, the idea I propose above (of running modules for separate randomly chosen subsample) is feasible only if the sample is large enough so that the sizes of the subsamples are adequate. Hence, a pre-condition for getting richer and more precise data from the existing surveys is to expand them both in scope and size.

Despite its appeal, redesigning existing surveys is not costless. In some cases, changing data collection methodologies for an existing data set is almost akin to start a new survey. I stress that this does not mean that modifying or enhancing existing surveys would be a bad idea. But if one is interested in studying dynamics of aggregate statistics (or individual dynamics across "old" and "redesigned" waves), then there will be some comparability problems that would need to be addressed.[19] One way to do this would be to have the new survey running in parallel with the old one for a few years, time and budget allowing.

4.2.Starting from scratch

There are pros and cons from initiating a brand new household survey. On the one hand, a new survey is a "blank slate" opportunity, where one can potentially fix the problems that plague existing surveys all at once (especially if they are well understood, which is however not totally obvious, i.e., under-reporting in the CEX). On the other hand, continuing existing surveys provides comparability of findings and data over time (especially if a panel component is available).

[19] For example, when the CPS was redesigned (in 1994) to improve its quality and expand its scope, researchers faced several issues related to comparability of labor market variables. There is a small literature discussing how data can be adjusted to ensure comparability over time of aggregate series.

I will discuss here briefly the advantages of starting a new survey and what this new survey should ideally include if the opportunity of running it arises. Of course, before making a definite deliberation it would be important to know the benefits and costs of conducting a new survey. Would the new survey result in discontinuing existing valuable surveys such as the PSID, for example? How large or detailed could the new survey be? How would a new survey compare with the "Gemini" redesigning project currently underway at the CEX? If there are areas of major overlap, does it make sense to duplicate efforts? It is also worth stressing that even if a new survey initially faced some problems, most of them will likely be fixed through a "learning-by-doing" process, in the same way as existing consumption survey have done over the years. This, of course, depends on the willingness of the institution running the survey to improve the data collection strategies. Historically, government-funded institutions like the BLS (perhaps because of the greater red tape involved) have been more reluctant to address problems with their surveys (the slow development of the CEX re-design being an example), while surveys administered by universities, research-oriented institutions, or other private institutions have made radical transformations of their surveys in a relatively short time (the case of PSID with the 1999 re-design) and obtained generally good results in terms of matching aggregate trends, etc.

4.2.1. What kind of information should the new survey contain?

Ideally, a new database would contain information helping researchers addressing most of the interesting topics that people are currently working on and that I summarized in Section 2, while striking a balance with costs (which typically affect the number and type of questions one may ask). This is a key issue of course. Recognizing the limitations of survey data collection, most of the valuable data sources used by economists and other social scientists typically come

with a precise "identity". For example, the CEX is primarily about household spending; the PSID (at least originally) was about poverty dynamics; the SIPP about welfare and social insurance use; the SCF about assets; the HRS focused on older people; the NLSY on younger people; and so forth. Would the new survey be primarily about household spending? And if so, how different would be from the CEX? Starting a new survey implies that there is a strong research program that backs it up. This does not mean that it is not a good idea to have new (and perhaps improved) sources of data. However, it is equally important to know what niche the new survey is designed to occupy, as there is a strong tradeoff between the time that the interview unit can devote to answer survey questions and the amount of detail one can collect on specific topic. For example, if the survey contains a lot of questions on financial wealth, it may be problematic having lots of details on networks, health, or subjective income data, etc.[20]

Assuming we are interested in collecting accurate spending data, some guidance is available from past recent experience. Crossley and Winter (2012) is an important reference surveying what have we learned regarding the optimal way to ask spending questions in survey data. They report the following general findings from the literature. First, the way in which the survey is conducted (telephone interviews, internet, leave-behind booklets, etc.) affects response behavior, but there is no obvious gold standard. As they write, there is "no systematic, controlled experimental study of how survey mode affects response quality in expenditure surveys". Second, as in the CEX, data can be collected by recall or diary strategies, each with its own issues. Recall data suffer from telescoping (attributing spending made in an earlier period to the

[20] Another possible framework for a new survey (which would also help giving it a precise identity) would be to interview a single cohort, such as all those born in the 1980s. This sampling scheme gives fruits only in the long term and cannot be compared with aggregate data. However, it has several strengths. For example, it can be used to study how network formation and evolution influence consumption dynamics or income shocks. It has the great advantage of an extreme large sample size for the chosen cohort.

current period), which grows in severity as the recall period increases. Diaries suffer from infrequency of purchases and declining reporting, induced either by fatigue (individuals stop bothering to report spending in the diary) or by a sort of Hawthorne effect (individuals realize they spend too much and cut back on it). As we have commented above, there is evidence that recall data do a better job than Diary data in replicating broad aggregate trends from NIPA. Some authors have concluded that annual recall may give good quality data as long as is used in conjunction with a cash-flow reconciliation or "balance edit".[21] Third, while asking more disaggregate consumption data gives higher estimates of the totals, there is a trade-off with accuracy, as increasing disaggregation is perceived as more intrusive and burdensome. The answer to the question of how disaggregated data should be is "it depends". Tests of lifecycle behavior (response of consumption to income shocks, say) would do well with less disaggregated data; while a study on the importance of home production, or one testing unitary vs. collective models of behavior, will probably require as much disaggregation of spending as possible. Fourth, it is useful to adopt different reference periods for different categories of goods. The PSID adopts this strategy (households can choose whether the reported amount they spend on food refers to a week, a month, etc.), while in the CEX it is fixed (monthly spending).

Keeping in mind these lessons, a new survey could contain a categorization of consumption goods into relatively broad aggregates, rather than attempting to replicate the level of details available in the CEX (this would allow to add questions on other topical themes). One option would be to have global questions on large aggregates (utilities, food, clothing, etc.) and then reserve special modules for a more in depth analysis of a specific category (i.e., clothing). If the

[21] The way this typically works is as follows. Households first report income and asset data. If their reported consumption differs too much from their reported income, the interviewer may ask additional questions to reconcile the income, asset and spending responses. In longitudinal data, interviewers may use previous income, asset holdings and spending as anchors for current reports.

CEX remains as is (and hence it is used to address the CPI weight construction issue), then this new survey does not need to include extreme details on consumers' spending, which induces interview fatigue and eventually may result in systematic measurement error and under-reporting. Survey experts' opinion may be solicited regarding the most efficient way to posit questions, and in which order, in order to maximize item response and data reporting quality. Having people providing electronic records, as I suggest below, would be ideal and at the same time reduce respondents' burden.

Other information should be included in order to link consumption behavior to socioeconomic characteristics. First, it is perhaps unthinkable to have an accurate view of consumption behavior without also collecting information on key covariates (or simultaneous endogenous variables) such as income (and its components, especially transfers), wealth (and its components, especially to look at the distinction between liquid and illiquid wealth, etc.), labor supply (and offered wage if not working), and detailed residence information (as geography seem to play an important role as recent work by Raj Chetty and co-authors, has demonstrated).

It would be also desirable to collect information on other domains, such as: (a) Time use information, in order to address home production-related questions; (b) information on private consumption vs. public good consumption (who does the expenditure refer to? Is it exclusive? etc.); (c) prices paid for relatively homogenous goods (i.e., a case of soda); (d) subjective questions aimed at measuring preferences (risk aversion, discount rate, etc.), income expectations and risk (as described in Manski, 2004), and marginal propensities to consume (as in Shapiro and Slemrod, 2003; Jappelli and Pistaferri, 2014), which in my personal experience can be fruitfully used in structural models of behavior, especially when coupled with panel data allowing researchers to "probe" the validity of subjective reports by confronting it with actual

behavior; and (e) network formation questions: identification of friends, relatives, co-workers as peers, etc. As suggested above, these questions could be part of special modules allocated to randomly selected subsamples, rather than being asked to all households in the sample. Alternatively, the new survey could be more focused and choose to collect detailed data on consumption/key covariates and one or two important domains: expectations and time use data, or expectations and social networks information, prices and intra-household consumption allocation, etc.

Given that the survey starts from scratch and has no history, it would be important to use life history questions to obviate the absence of a backward panel (marriage history, residence history, employment history, while the age of kids could help reconstruct fertility history). However, these questions are extremely expensive in terms of time spent by the respondent (especially for individuals later in their life cycle). In fact, in the SIPP such questions are relegated to special topical modules.

Having respondents providing social security numbers would allow researchers to also recover the history of income records by matching the data with IRS, SSA, or other registries – with little to no burden for the respondents. This is a key point. "Big data" means frictionless access to data without any commitment on the part of the respondent besides her initial willingness to participate. Needless to say, confidentiality issues would be non-negligible. The new survey should avoid use proxy responses which notoriously reduce the quality of the data; moreover, it may make the collection of "individual" consumption information complicated.

The disadvantage of starting a new survey is that, as discussed earlier, in the US survey non-response is increasing. Indeed, it is becoming increasingly difficult to reach households for survey purposes. Partly this reflects an increased value of time and increased labor force

participation of women (which makes it hard to find respondents at home during business hours); partly this reflects the fact that traditional ways of reaching households have become obsolete – some households do not have a landline telephone, or have technologies that screen out calls from data collection agencies. There may be increased response rates from Internet-based surveys, but the representativeness of such surveys (made up of *professional* respondents) may raise concerns.

4.3.Data collection strategies for the XXI century?

A third route I want to discuss briefly is the collection of administrative data on consumption. There is currently a "Big Data" revolution taking place in many fields, including economics. Researchers are increasingly using complex, large, and sophisticated administrative databases to answer important questions in Labor Economics, Public Finance, International trade, etc. For example, people have managed to access micro-level IRS data, Social Security records and (mostly in Europe) matched employer-employee information. Despite its quantitative relevance, consumption is lagging behind in administrative data collection effort and study. Four exceptions can be cited, and they heed important lessons for future efforts. One is spending data available from credit card companies (as in Gross and Souleles, 2002); another is the use of longitudinal administrative tax record information on income and wealth to create consumption starting from the budget identity (i.e., $C=Y-\Delta A$ as in Browning and Leth-Petersen, 2003, and De Giorgi, Frederikssen and Pistaferri, 2014 for Denmark, Koijen, Van Nieuwerburghy, and Vestmanz, 2013, for Sweden); data on spending, income, assets, etc., for consumers using online financial aggregators such as Mint.com or Check.com (as in Baker, 2014, and Gelman et al., 2013, respectively); finally, scanner data information coming from Nielsen Homescan data sets as in

Handbury (2014) and Broda and Weinstein (2008). Scanner data are the closest equivalent to administrative data we can get on spending. However, it should be noted that the data refer primarily to food/grocery store items. A consortium of Stanford and Berkeley economists were recently able to gain access to detailed spending data provided from one of the largest national grocery chains in the US (these data were used by Mas and Moretti, 2009, and Einav, Leibtag and Nevo, 2010, among many). But once more, data are limited to supermarket items and naturally there are several issues regarding representativeness of the sample. While these are remarkable steps ahead, the data sources above are either not representative of all households or not representative of all the goods that people buy.

Recalling expenditures is notoriously difficult. It can be achieved accurately with multiple probing questions, but this may have the effect of increasing survey item non-response or inducing survey fatigue.[22] As an alternative, consumers could allow interviewers to access electronic records containing details on what they buy. There is already an ongoing discussion on how these data collection methodologies should be introduced. One potentially fruitfully way is to use internet financial aggregators. Once a household has been sampled, one way to collect detailed reliable information on overall spending, income, etc., is to ask the household to open a mint.com (or check.com) account, if they do not already have one. Mint.com is a web-based personal financial management service that aggregates, in one single platform, all checking accounts, saving accounts, credit card accounts, investment accounts, and more (such as the value of one's house, through Zillow.com). If respondents agree to post all their accounts onto the platform, spending gets recorded automatically. Moreover, an important by-product of the data set is that for consumers who receive their pay-check through direct deposit, one also

[22] One way to reduce survey fatigue is to allocate different households to different spending categories, but this can only work when we use pseudo-panels or cohort data.

observes the identity of the employer, opening up opportunities for addressing a variety of research questions (such as consumption network effects of the type addressed by De Giorgi, Frederikssen, and Pistaferri, 2014; *consumption insurance* within the firm, as in Guiso, Schivardi and Pistaferri, 2008; tests of "home bias" effects, and so forth). A follow-up interview can be used to allow interviewer to inquire about specific items, what appear as duplicates, etc., and to inquire about cash income and cash consumption. Providing consumers with internet tablets could help increase their involvement with the survey.

Needless to say, this proposal is ambitious. First, it assumes that all people have Internet access. While this is still not universal in the U.S., things are changing extremely rapidly. According to a Census Bureau (2013) publication, in 1997 little less than 37% of US households reported having a computer at home, and about 18% reported accessing the Internet. By 2011, the two figures had jumped to 76% and 72%, respectively. This seems to suggests that it is not unconceivable that within the next decade, computer ownership and Internet usage would be as ubiquitous as ownership of a fridge or a TV set are today (in 2009, 99.9% and 98.7% of all US households owned these two durable goods).[23]

The other assumption is that people would be willing to record all expenditures on the website, but this actually is the least of the problems. In fact, the most useful aspect of accessing data from personal financial websites such as Mint.com or Check.com is that participants record *no* information on their own (i.e., there is no involvement on the part of the consumer). Once the user authorizes the website to add a given account, information on account usage (spending, income, etc.) is flawlessly and automatically uploaded. This reduces dramatically the extent of measurement error and recall biases that plague survey data. Moreover, Mint.com data permit the

[23] According to the 2009 Residential Energy Consumption Survey.

construction of a broad spending aggregate almost effortlessly. It uses the payee's/merchant's name/label (i.e., Safeway in the screenshot below) to assign purchases to categories (auto and transports, bills and utilities, etc.). The software is quite accurate. For example, in Fig. 7, Mint.com appears to accurately categorize all purchases with (somewhat understandably) the exception of those made during a business trip in a foreign country (the UK).

[Fig. 7 here]

The users can volunteer information if the software fails to identify the spending category (as in the UNCATEGORIZED cases in bold in the Figure) or when the software misallocates the purchase.

The final, and most critical assumption, is that people need to be willing to share the extreme details collected by Mint.com with independent researchers. This is far from obvious and probably the largest obstacle for a more widespread use of these data. At the moment, the extreme details contained in financial aggregator data are available only to a few researchers (and the identity of users appropriately masked). If the objective is to obtain highly reliable and disaggregated consumption data, one option would be to partner with the financial software companies in the same way as Baker (2014) has done – but this would mean giving up on the opportunity of collecting information that are not contained on the site (such as expectations or more detailed demographics and socio-economic variables).[24] [25] There are important concerns regarding the representativeness of samples drawn from the personal finance software, but this is

[24] While this is not ideal, it should be noted that the large scale administrative data used by labor and public finance economists (IRS data or employer-employee data drawn from social security or tax record data) also typically lack important information, such as education or hours of work.

[25] It should also be noted that the software companies may have little incentives to share the data (especially if widespread usage of the data may be interpreted by some users as a breach of the confidentiality agreement) or ask users to answer survey-type questions. Alternatively, a new survey could start with a platform *similar* to Mint.com, but managed by the agency running the survey.

easing over time as the websites become more popular and trust is built about their security. For example, Baker (2014) reports that the website whose data he is using grew ten-fold (from 300,000 to 3 millions) between 2007 and 2012. In keeping with this growth, the sample has become more representative of the US general population (see his Table 1). The typical user is still more likely to be male, married, homeowner, and richer than the average US individual, but the differences are shrinking. As for consumption, Baker (2014) compares his data with those from the Retailer Census for different goods and finds that most trends are well replicated, with the exception of spending on vehicles, since individual purchases may differ in important ways from receipts recorded by the sellers. Overall, there seems to be a lot of research potential from the use of these data.

6. CONCLUSIONS

Access to accurate microeconomic data on consumption is important to be able to answer a wide range of questions of academic, policy, and welfare relevance. Despite this observation, data on household consumption and spending in the US are few and problematic. On the first point, it is remarkable that there is practically only one data set with a comprehensive measure of consumption (CEX), while there are plenty of data sets containing information on individual labor earnings, say (CPS, SIPP, PSID, CEX, ACS to cite a few). On the second point, existing survey data such as the CEX or the PSID are valuable sources of information but clearly far from perfect. For these reasons, it is not surprising that calls for improving existing survey data or complement them with newly collected survey data have intensified lately.

In my personal view, a brand new survey, *per se*, is unlikely to solve the problems that plague the existing surveys, especially if it ends up being a mere replica of what is already

available. Indeed, it would be much more valuable to invest in improving the existing surveys along the lines discussed in this essay (and beyond). In contrast, I would be much more supportive of a new survey that is fundamentally novel in its data collection approach (i.e., it mixes conventional strategies with more administrative-type methods). Cleary, some of the ideas discussed in the essay (access to administrative measures of consumption, or matching consumption survey data with social security or tax records data) are hard to implement in practice and may run against important privacy and confidentiality issues.[26] However, its feasibility (or feasibility of variants thereof) could at least be assessed by running experimental, small pilots. For example, the NAS panel in charge with proposing the re-design of the CEX also recommended exploring the idea of providing households with tablets to record purchases, but this has not been tried on a probability sample yet, nor done in a panel over a long time frame. A feasibility study should be attempted to understand whether this is a viable option and how reliable the data so collected are. The "big data" revolution offers many challenges but also countless research opportunities, which should be exploited in the consumption research area exactly as is currently done in other economics fields like IO, Labor and Public Finance. Going forward, it would be crucial to understand how to reconcile the difficulties involved in collecting quasi-administrative data on spending with sample representativeness and privacy concerns.

[26] For example, Obtaining SSNs from respondents raises important concerns about consent rates. The more these kinds of burdens are imposed on survey respondents, the more selective the survey is likely to become.

REFERENCES

[1] Adams, Williams, Liran Einav, and Jonathan Levin (2009), "Liquidity Constraints and Imperfect Information in Subprime Lending", American Economic Review, 99:1, 49–84

[2] Aguiar, Mark A. and Mark Bils (2011), "Has Consumption Inequality Mirrored Income Inequality?," NBER Working Papers 16807.

[3] Aguiar, Mark, and Erik Hurst (2007), "Lifecycle Prices and Production", American Economic Review, 97(5), 1533-59.

[4] Attanasio, P. Orazio, Pinelopi Koujianou Goldberg, and Kyriazidou, Ekaterini (2008), "Credit constraints in the market for consumer durables: Evidence from micro data on car loans", International Economic Review, Vol. 49, No. 2., pp. 401-436

[5] Attanasio Orazio, Erik Hurst, and Luigi Pistaferri (2014), "The Evolution of Income, Consumption, and Leisure Inequality in the US, 1980-2010", in Improving Measurement of Consumer Expenditures (eds. Chris Carroll, Thomas Crossley, and John Sabelhaus)

[6] Attanasio Orazio, and Luigi Pistaferri (2014), "Consumption inequality over the last half century. Evidence using the new PSID consumption measure", American Economic Review P&P (forthcoming).

[7] Baker, Scott (2014), "Debt and the Consumption Response to Household Income Shocks", mimeo.

[8] Bee, Adam, Bruce D. Meyer and James X. Sullivan (2014), "The Validity of Consumption Data: Are the Consumer Expenditure Interview and Diary Surveys Informative?,", in Improving Measurement of Consumer Expenditures (eds. Chris Carroll, Thomas Crossley, and John Sabelhaus).

[9] Bertrand, Marianne and Adair Morse (2013), "Trickle down consumption", NBER WP 18883.

[10] Bloom, Nick (2014), "Fluctuations in Uncertainty", Journal of Economic Perspectives.

[11] Blundell, Richard, Luigi Pistaferri and Itay Saporta-Eksten (2012), "Consumption Inequality and Family Labor Supply", NBER Working Papers 18445.

[12] Broda, Christian and David Weinstein (2008), "Prices, Poverty and Inequality", AEI Well-being Series Monograph.

[13] Browning Martin, Mette Ejrnaes and Javier Alvarez (2010), "Modelling Income Processes with Lots of Heterogeneity," Review of Economic Studies, Oxford University Press, vol. 77(4), pages 1353-1381.

[14] Browning, Martin, and Leth-Petersen, Soren, (2003), "Imputing Consumption from Income and Wealth Information." The Economic Journal, Vol. 113, No. 488, pp. F282-F301

[15] Browning, Martin and Annamaria Lusardi (1996), "Household Saving: Micro Theories and Micro Facts," Journal of Economic Literature, 34(4), pages 1797-1855.

[16] Cherchye, Luc, Brian De Rock and Frederic Vermeulen (2012), "Married with children. A collective labor supply model with detailed time use and intrahousehold expenditure information", American Economic Review, 102, 3377-3405.

[17] Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014), "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", NBER Working Paper # 19843.

[18] Census Bureau (2013), "Computer and Internet Use in the United States".

[19] Coibion Olivier, Yuri Gorodnichenko, Lorenz Kueng, and Juan Silvia (2012), "Innocent Bystanders? Monetary Policy and Inequality in the U.S.", mimeo.
[20] Crossley, Thomas F. and Joachim K. Winter (2012), "Asking Households about Expenditures: What Have We Learned?," NBER Chapters, in: Improving the Measurement of Consumer Expenditures National Bureau of Economic Research, Inc.
[21] DeBacker, Jason, Bradley Heim, Vasia Panousi, Shanthi Ramnath and Ivan Vidangos (2013), "Rising Inequality: Transitory or Persistent? New Evidence from a Panel of U.S. Tax Returns," Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 46(1 (Spring), pages 67-142.
[22] De Giorgi Giacomo, Anders Frederikssen and Luigi Pistaferri (2014), "Consumption network effects", mimeo.
[23] Della Vigna, Stefano (2009), "Psychology and Economics: Evidence from the Field", Journal of Economic Literature, Vol. 47, pp. 315-372.
[24] Della Vigna, Stefano, and Joshua Pollet (2007), "Demographics and Industry Returns", American Economic Review, 97, pp. 1167-1702.
[25] Dominitz, Jeff, and Charles F. Manski (1997), "Using expectations data to study subjective income expectations," Journal of the American Statistical Association 92, 855-67.
[26] Einav, Liran, Ephraim Leibtag and Aviv Nevo (2010), "Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?," Quantitative Marketing and Economics, Springer, vol. 8(2), pages 207-239, June.
[27] Gelman, Michael, Shachar Kariv, Mattew Shapiro, Dan Silverman, and Steven Tadelis (2013), "Harnessing Naturally-Occurring Data to Measure Income and Spending Accurately in Real Time", mimeo.
[28] Gerardi, K., H. Rosen and P. Willen (2010), "The Impact of Deregulation and Financial Innovation on Consumers: The Case of the Mortgage Market", The Journal of Finance 65(1).
[29] Gottschalk, Peter and Robert Moffitt (1994), "The growth of earnings instability in the US labor market", Brookings Papers on Economic Activity, 217-272.
[30] Gross, David B. and Nicholas S. Souleles (2002), "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data", Quarterly Journal of Economics, 117(1).
[31] Guiso, Luigi, Fabiano Schivardi and Luigi Pistaferri (2008), "Insurance within the firm", Journal of Political Economy.
[32] Guvenen, Fatih, Serdar Ozkan and Jae Song (2014), "The Nature of Countercyclical Income Risk", forthcoming, Journal of Political Economy.
[33] Handbury Jessie (2014), "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities", mimeo.
[34] Hastings, Justine, and Ebonya Washington (2010), "The First of the Month Effect: Consumer Behavior and Store Responses." American Economic Journal: Economic Policy, 2(2): 142-62.
[35] Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante (2010), "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," Review of Economic Dynamics, vol. 13(1), pages 15-51, January.
[36] Hendal, Igal, and Aviv Nevo (2013), "Intertemporal Price Discrimination in Storable Goods Markets", American Economic Review (forthcoming).

[37] Jappelli Tullio, (1990), “Who is credit constrained in the US economy?”, Quarterly Journal of Economics, vol. 105, n. 420, 219-234.
[38] Jappelli Tullio, and Luigi Pistaferri (2010), "The consumption response to income changes", Annual Review of Economics 479-506.
[39] Jappelli Tullio, and Luigi Pistaferri (2014), "Fiscal policy and MPC heterogeneity", American Economic Journal: Macroeconomics (forthcoming).
[40] Johnson David, Parker Jonathan, Nicholas Souleles (2006), “Household expenditure and the income tax rebates of 2001”, American Economic Review.
[41] Kaplan, Greg, and Guido Menzio (2013), “Shopping Externalities and Self-Fulfilling Unemployment Fluctuations”, mimeo.
[42] Kaplan, Greg, and Gianluca Violante (2013), “A Model of the Consumption Response to Fiscal Stimulus Payments”, mimeo.
[43] Koijen, Ralph, Stijn Van Nieuwerburghy, and Roine Vestmanz (2013), “Judging the Quality of Survey Data by Comparison with "Truth" as Measured By Administrative Records: Evidence from Sweden”, mimeo.
[44] Krueger, Dirk and Fabrizio Perri (2006), "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," Review of Economic Studies, 73(1): 163-193.
[45] Lucas, Robert (1976). "Econometric Policy Evaluation: A Critique". In Brunner, K.; Meltzer, A. The Phillips Curve and Labor Markets. Carnegie-Rochester Conference Series on Public Policy 1. New York: American Elsevier. pp. 19–46
[46] Manski, Charles F. (2004), "Measuring Expectations," Econometrica, vol. 72(5), p. 1329-1376.
[47] Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." American Economic Review, 99(1): 112-45.
[48] Meyer, Bruce and James Sullivan (2009), "Five Decades of Consumption and Income Poverty," NBER Working Papers 14827.
[49] Mian, Atif, K. Rao, and Amir Sufi (2013), “Household Balance Sheets, Consumption, and the Economic Slump”, Quarterly Journal of Economics, forthcoming
[50] Misra, Kanishka and Paolo Surico. 2012. “Consumption, Income Changes and Heterogeneity: Evidence from Two Fiscal Stimulus Programmes”. London Business School, mimeo.
[51] Padula Mario, 1999. "Euler Equations and Durable Goods," CSEF Working Papers 30, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.
[52] Passero, William, Thesia I. Garner, and Clinton McCully (2014), “Understanding the Relationship: CE Survey and PCE”, in Improving the Measurement of Consumer Expenditures, Christopher Carroll, Thomas Crossley, and John Sabelhaus, editors.
[53] Pistaferri, Luigi (2001), "Superior Information, Income Shocks, And The Permanent Income Hypothesis," The Review of Economics and Statistics, MIT Press, vol. 83(3), pages 465-476, August.
[54] Rajan, Raghuram (2010), Fault Lines, Princeton University Press.
[55] Sahm, Claudia R., Matthew D. Shapiro, and Joel Slemrod (2010). “Household Response to the 2008 Tax Rebates: Survey Evidence and Aggregate Implications,” in Brown, Jeffrey R. ed., Tax Policy and the Economy, Vol. 24. National Bureau of Economic Research, 69-110.
[56] Shapiro Matthew D., Slemrod Joel. 1995. “Consumer Response to the Timing of Income: Evidence from a Change in Tax Withholding”, American Economic Review 85, 274-283.

[57] Shapiro Matthew D., Slemrod Joel. 2003. “Did the 2001 Tax Rebate Stimulate Spending? Evidence from Taxpayer Surveys”, in Tax Policy and the Economy. Vol. 17, ed. James M. Poterba, Cambridge, MA: MIT Press.
[58] Souleles, Nicholas S., Sumit Agarwal, Chunlin Liu (2007), The Reaction of Consumer Spending and Debt to Tax Rebates – Evidence from Consumer Credit Data, Journal of Political Economy, 115(6).
[59] Wolff, Edward N., Lindsay A. Owens, and Esra Burak (2011), "How Much Wealth Was Destroyed in the Great Recession?". The Great Recession. Eds. David Grusky, Bruce Western and Christopher Wimer. New York: Russell Sage,127-158.

Fig. 1: A taxonomy of consumption response to changes in economic resources

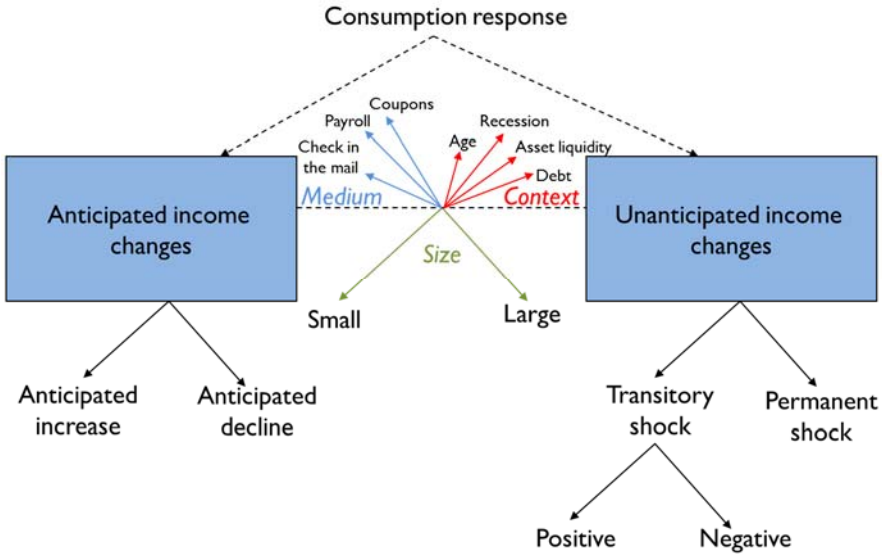

Fig. 2: Consumption growth rates across the Great recession periods for the bottom and top income quintiles.

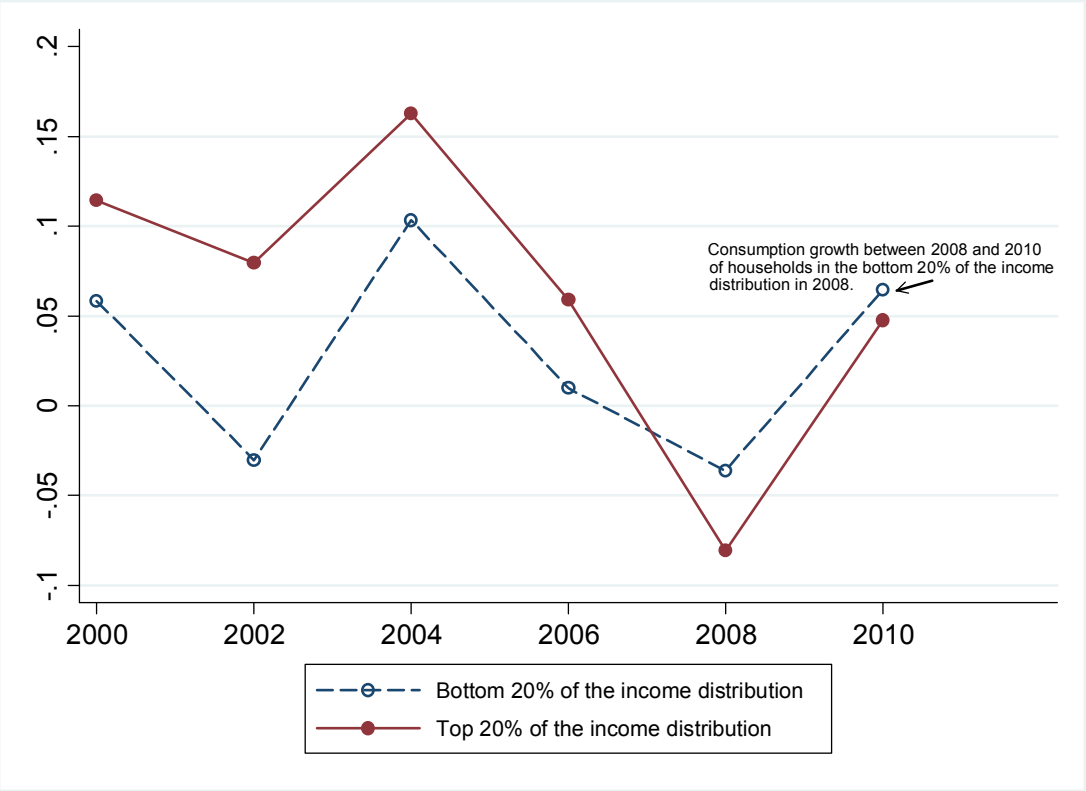

Fig. 3: The evolution of consumption inequality over the life cycle for cohorts born in the 1940s, 1950s, 1960s, and 1970s.

Fig. 4: Consumption Mobility in the US, 2000-2010

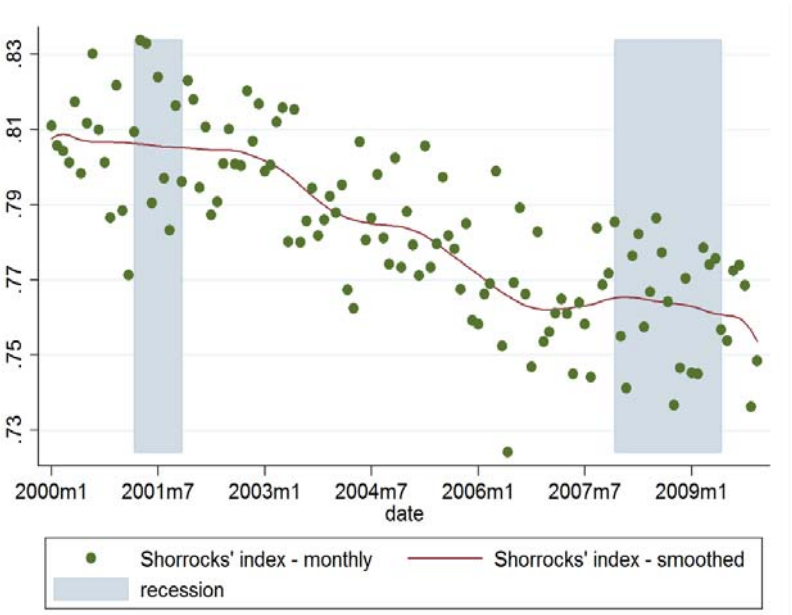

Fig. 5: Comparing the CEX and NIPA

(Source: Heathcote, Perri and Violante, 2010)

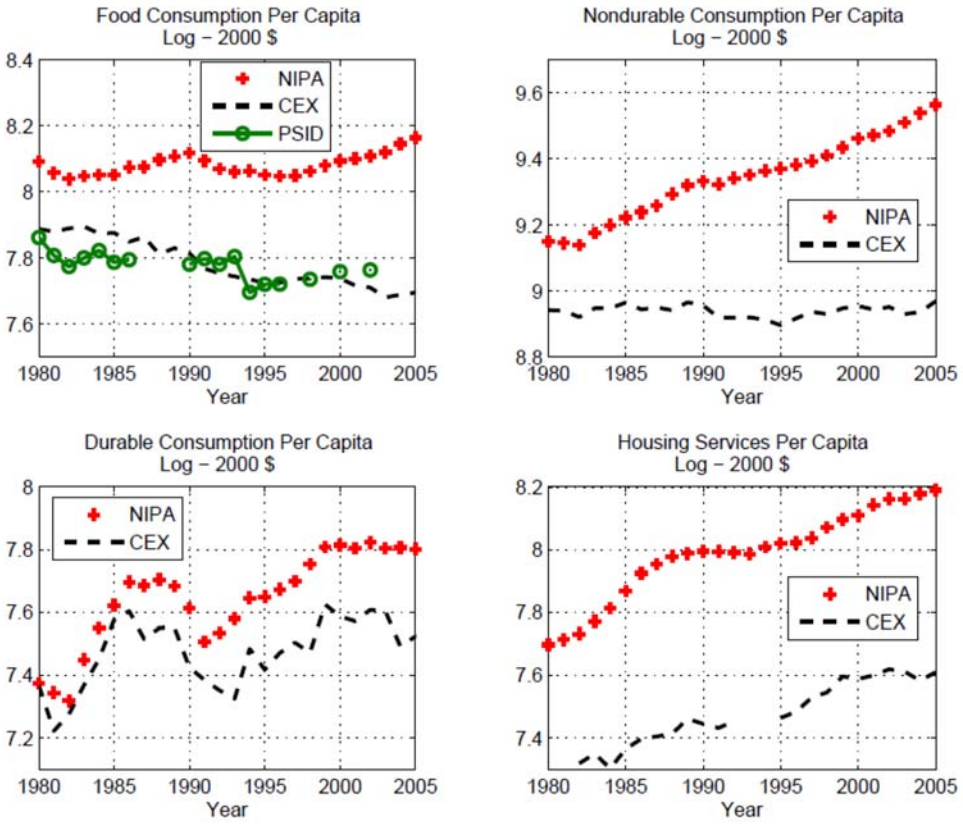

Fig. 6: Growth rates in the redesigned PSID and NIPA

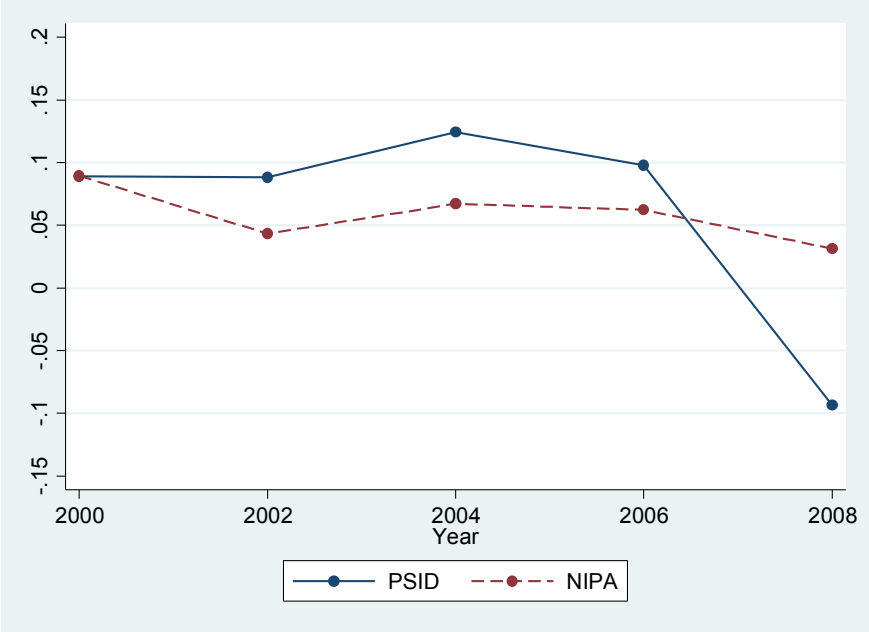

Fig. 7: An example of a Mint.com screenshot

	Date	Description	Category		Amount
	10/14/10	Safeway	Groceries		–$28.24
	10/15/10	Union 76 EDIT DETAILS	Gas & Fuel		–$41.39
	10/15/10	Palo Alto Medical Clinic	Doctor		–$35.00
	10/16/10	Sierra Fountain Svc	Home		–$9.00
	10/17/10	Chevron	Gas & Fuel		–$40.41
	10/18/10	Safeway	Groceries		–$37.71
	10/18/10	Whole Foods	Groceries		–$37.43
	10/18/10	Valero	Gas & Fuel		–$30.45
	10/19/10	Macy's	Clothing		–$27.30
	10/19/10	Caffe Del Mond	**UNCATEGORIZED**		–$4.37
	10/19/10	Apple	Electronics & So...		–$20.76
	10/19/10	Supershuttle	Auto & Transport		–$30.68
	10/19/10	Macy's	Clothing		–$87.05
	10/19/10	Stanford Alumni Cafe	Restaurants		–$4.42
	10/19/10	Gateway	Electronics & So...		–$8.00
	10/20/10	Heathrow Rail Link	**UNCATEGORIZED**		–$28.49
	10/20/10	Sainsburys Smkt Paddington	**UNCATEGORIZED**		–$11.17
	10/20/10	Delta	Air Travel		–$862.31