the review of income and wealth

Review of Income and Wealth
Series 55, Number 2, June 2009

WHERE IS THE MISSING CREDIT CARD DEBT? CLUES AND IMPLICATIONS

by Jonathan Zinman*

Dartmouth College

I create comparable estimates of aggregate credit card use based on household data from the Survey of Consumer Finances (SCF) and industry data. The two sources match up well on credit card charges and fairly well on account totals. But the SCF always yields much lower estimates of revolving debt. My estimated lower bound for the discrepancy in 2004 is half of the revolving credit card debt total implied by industry data. There is no obvious source for this remaining discrepancy and some evidence that the discrepancy has grown over time. Such growth is worrisome because it parallels substantial changes in credit card use and in the pool of credit card users, suggesting that the discrepancy could be driven by household underreporting that is correlated with unobserved heterogeneity. This correlation could confound inference on the relationship between credit card borrowing and outcomes of interest like household financial condition, consumption paths, and portfolio choice. Given this possibility it is critical to continue developing evidence on whether and why household surveys undercount credit card borrowing.

1. Introduction

A casual comparison of United States data suggests a big discrepancy between the balance sheets of credit card borrowers and lenders. Lenders report owning three times as much consumer credit card debt in the Federal Reserve Board's *Statistical Release G.19: Consumer Credit* (the "G.19") as households report owing in the Federal Reserve Board's *Survey of Consumer Finances* (SCF). In 2004 the discrepancy was $537 billion.

Understanding the reasons for this "missing" credit card debt is critical for studies of household finance such as intertemporal and portfolio choice and of credit market (in)efficiency. Credit cards are the largest source of non-mortgage borrowing in the U.S. and are rapidly approaching that status elsewhere. Credit cards are also plausibly the *marginal* source of borrowing and consumption for most U.S. households.[1] Researchers have increasingly turned to financial micro-data from household surveys like the SCF for several reasons. One is practical: administrative data is often hard to obtain. Other reasons are substantive. The

Note: I thank the Payment Card Center at the Federal Reserve Bank of Philadelphia for hosting me as a Visiting Scholar, Adeel Ikram for research assistance, and Mark Furletti, Kathleen Johnson, Art Kennickell, Bill Lang, Doug Staiger, the editor Marcel Timmer, two anonymous referees, and especially Christopher Ody for comments.

*Correspondence to: Jonathan Zinman, Department of Economics, Dartmouth College, 314 Rockefeller Hall, Hanover, NH 03755, USA (jzinman@dartmouth.edu).

[1]The 2004 Survey of Consumer Finances finds that 72 percent of U.S. households held a general purpose credit card; in contrast, only 12 percent had a home equity line of credit. Average utilization of credit card credit limits was only around 20 percent (http://www.nationalscoreindex.com/NSI_Site/USScore.aspx; Nilson #705, 1999). See Sullivan (2008) for evidence on consumption smoothing with credit cards.

SCF is "the most comprehensive survey on household wealth" (Antoniewicz *et al.*, 2005, p. 7). Households hold credit cards (and other balance sheet items) from several different issuers, and hence a household survey is often the only way to get a comprehensive picture of the household balance sheet. Household surveys may also be better suited than aggregates or administrative microdata for collecting information used to study heterogeneity in financial outcomes and model parameters.

While household survey data on credit card borrowing is useful in principle, the limited available evidence suggests a potentially critical practical limitation: severe undercounting. Two double-blind validation studies on other forms of unsecured consumer borrowing show that more than half of borrowers, known from lender-side data to have borrowed recently, do not report *any* recent borrowing. Elliehausen and Lawrence (2001) show underreporting of payday loan borrowing in phone surveys. Karlan and Zinman (2008) find underreporting of a payday-like loan product in a South African in-person survey.

Researchers using the SCF typically ignore the possibility of underreporting (or of undercounting more generally), or they scale up SCF borrowing with an average undercounting factor that matches total, weighted-up borrowing in the SCF to the G.19 aggregate. The G.19 is widely considered to be accurate *a priori* (because lenders have strong incentives to report truthfully), and has been validated by a third party (Furletti and Ody, 2006).[2] But for most applications these approaches to SCF undercounting will permit valid inference if and only if two conditions hold. First, one must use the correct scalar for average undercounting. Second, an adjusted average undercounting factor is only broadly useful if household (under)counting is homogenous or effectively so; e.g. if everyone underreports by the same factor, or if underreporting is not correlated with parameters and outcomes of interest.

Unobserved heterogeneity in respondent reporting of credit card debt could present a serious problem for researchers across a range of literatures and methods. It would confound attempts to estimate the distribution, incidence, and determinants of borrowing costs; this concern applies, for example, to the literature exploring why households simultaneously borrow expensively on cards while lending in low-yielding demand deposits. Unobserved heterogeneity in reporting might also bias estimates of parameters or model fit; these concerns apply, for example, to literatures estimating the effect of debt burden on financial distress, the effect of house prices on consumption, and the fit of alternative models of intertemporal and portfolio choice. I discuss implications of reporting heterogeneity more in-depth in Section 5. The challenge in assessing whether unobserved

[2]The main source for the G.19 revolving credit estimate, the quarterly Report of Condition and Income ("Call Reports"), is widely used by researchers and considered to be accurate. Commercial banks (who issue and own over 80 percent of credit card outstandings) are closely scrutinized by regulators; this supervision presumably constrains the ability of issuers to misreport on a scale large enough to explain the wedge between the G.19 and the SCF. More to the point there is no *a priori* reason why, after adjusting for the definitional differences as I do below, one would expect issuers to *over*report. If anything the relatively steep capital charge on credit card assets would push banks to *under*report. Chargeoffs or delinquencies could explain a small fraction of the wedge (they are equal to about 4 percent of outstandings on average) if borrowers do not report their bad/late debt and issuers do.

heterogeneity is actually a problem is that measuring household reporting behavior is difficult without direct validation from lender data.

In this paper I explore whether unobserved heterogeneity is a problem using indirect methods feasible given the available data. The first finding is that the discrepancy between SCF and industry measures of aggregate credit card borrowing can not be explained by obvious definitional or coverage differences. The G.19 measure of *outstandings* includes several types of credit card use that are excluded from the SCF measures of credit card *revolving* by design. These include float (as other researchers have recognized),[3] business use of personal cards, and non-credit card lines of credit. After adjusting for several differences my estimated *lower* bound on the discrepancy for 2004 is $200 billion. This suggests that the SCF misses about half of revolving credit card debt.

A second finding is some evidence that the definition/coverage-adjusted debt discrepancy has been rising over time (e.g. my preferred estimates suggest that the SCF undercounted by a factor of 1.3 or 2.0 in 1989, and by 1.9 or 2.4 in 2004). This is worrisome given that credit card use and the pool of credit card users was also changing over this time period. Use shifted toward general purpose cards, and 26 million households (generally less creditworthy ones) entered the market.[4] This suggests that any SCF undercounting might be driven by changes in unobserved household characteristics that could confound inference.[5]

A third finding is that in contrast to debt, the SCF credit card charges and account totals match up relatively well with industry sources. This could be due to differences between industry sources (here Nilson Reports instead of the G.19) rather than anything about SCF. But if charges and accounts do ultimately prove to be reported more completely than debt it will limit the set of likely explanations for SCF undercounting of credit card debt.

The rest of the paper proceeds as follows. Section 2 describes the strengths and weaknesses of the data sources I use to measure credit card activity. Section 3 provides estimates of the discrepancy between SCF and industry sources that attempt to adjust for definitional and coverage differences. Section 4 discusses possible sources of the discrepancy. Section 5 explores whether the remaining discrepancy makes the SCF credit card borrowing data too problematic for broad use by researchers. Section 6 concludes.

2. Data Sources and Study Period

This section provides an overview description of my three main data sources for measuring credit card account holding, charging, and borrowing. Section 3 provides details on definitional and coverage differences, and on the related adjustments I make to facilitate comparisons.

[3]See, e.g. Gross and Souleles (2002), Johnson (2007), and Laibson *et al.* (2009).

[4]See Edelberg (2006) and Calem *et al*. (2006) for details on industry innovations that drove this credit expansion.

[5]For example, Table 1 suggests that the marginal credit card holders were poorer and more likely to be female than inframarginal ones. Karlan and Zinman (2008) find evidence suggesting that females and relatively low-income borrowers are more likely to underreport their expensive, unsecured debt.

TABLE 1

CHARACTERISTICS OF SCF MARGINAL CARDHOLDERS (I.E. HOUSEHOLDS WITH 1 GENERAL PURPOSE CARD)

Year	Income < Median	Education < Median	Age < 35	Nonwhite	Female	Unmarried	Renter	Job Tenure < Median	Home Tenure < Median
1989	0.61	0.49	0.21	0.14	0.26	0.39	0.29	0.55	0.46
1992	0.63	0.50	0.23	0.16	0.26	0.42	0.32	0.53	0.46
1995	0.65	0.56	0.22	0.18	0.29	0.47	0.35	0.56	0.50
1998	0.62	0.53	0.21	0.18	0.26	0.47	0.35	0.54	0.48
2001	0.67	0.54	0.21	0.19	0.30	0.49	0.34	0.54	0.45
2004	0.68	0.51	0.23	0.23	0.32	0.50	0.31	0.59	0.50
p-value on test of whether 2004 mean > 1989 mean	0.01	0.55	0.40	0.00	0.02	0.00	0.43	0.14	0.27

Notes: Proportions weighted with SCF variable x42001. Education is years completed. Education, age, race, gender, job tenure with current employer, and marital status are those of the survey respondents.

The primary industry source on credit card debt is the *Federal Reserve Statistical Release G.19: Consumer Credit* (G.19). The G.19 takes a snapshot estimate of outstandings (amounts owed to credit card issuers) on all consumer credit card accounts at month-end, making the appropriate adjustments so that securitized receivables do not get double-counted. G.19 credit card outstandings are comprised (though not easily disaggregated into) three categories of cards: those that can be used at many different merchants (general purpose cards),[6] those that can be used only at a particular store or chain (store cards), and those than can be used at particular fueling stations (gas cards). The G.19's revolving "total" also includes a small amount of non-credit card lines of credit, as detailed in Section 3.4.

Although one needs to be cautious about using administrative data for statistical research (see, e.g. Brackstone, 1987), both theory and evidence suggest that the G.19 produces an accurate estimate of outstanding credit card debt.

A priori, it seems likely that most credit card issuers have incentives to report accurately. The main source data for the G.19, the quarterly Report of Condition and Income ("Call Reports"), is scrutinized by bank safety and soundness regulators (commercial banks issue and own more than 80 percent of credit card outstandings).[7] This supervision presumably constrains the ability of issuers to misreport on a scale large enough to explain a large discrepancy between the G.19 and the SCF. More to the point there is no *a priori* reason why, after adjusting for the definitional differences as I do below, one would expect issuers to *over*report. If anything the relatively steep capital charge on credit card assets would push banks to *under*report. Chargeoffs or delinquencies could explain a small fraction of the discrepancy (they are equal to about 4 percent of outstandings on average) if borrowers do not report their bad/late debt and issuers do.

[6]General purpose cards are offered by the Visa, MasterCard, Discover, and American Express brands/networks, and most have revolving credit features. A small fraction of general purpose card activity is on "charge cards" (issued principally by American Express) that do not have revolving features—each month's balance must be paid in full.

[7]The Call Reports are also widely used by researchers; to my knowledge no one has uncovered serious reporting biases.

In practice, Furletti and Ody (2006) detail and check the sampling procedures used to construct the G.19 monthly revolving credit estimate. Although they suggest some minor modifications overall they find the G.19 to be "highly accurate."

But the G.19 does not provide any information on credit card charge volume or account tallies, so I combed various issues of *The Nilson Report* (Nilson) for issuer-side data on these margins. Unlike the G.19 and SCF, Nilson data varies in content and format across years. Consequently I reviewed every issue in the year following the comparison years in my sample (e.g. 2005 issues for 2004 data) for data that might be informative. Nilson does not provide details on its sampling procedures and hence its accuracy has not been validated (to my knowledge).[8]

The Survey of Consumer Finances (SCF) is the household-level data source used in my comparisons. It is the most comprehensive nationally representative source of data on credit card use and household finance more generally.[9] The SCF is conducted every three years and surveys around 4,000 households each wave. I focus on the 1989–2004 SCFs because earlier G.19s did not capture large swaths of outstandings.[10] There is no panel component to the SCF during this period. For each type of credit card use I construct aggregate estimates by inflating each household by its SCF-assigned population weight (variable x42001).[11]

The SCF questions and prompts are nearly identical across the six different surveys during my study period. The survey collects the data of interest starting with a question on whether anyone in the household has any credit cards or charge cards (with a prompt to distinguish these from debit cards). For respondents that answer "yes," the surveyor then asks five yes/no questions about whether anyone in the household has any of five different types of cards (general purpose revolving, store, gas, general purpose charge, other). Following each "yes" answer respondents are asked how many *accounts* they have of that type, and are specifically instructed to "not count duplicate cards on the same account or any business or company accounts." Respondents reporting one or more accounts for a given card type are then asked: "On your last bill(s), roughly how much were the new charges made to this/these account(s)?" and "After the last payments were made on this/these account(s), roughly what was the balance still owed on these accounts?" These produce account type-level measures of what I label "recent charges" and "revolving debt."

[8]The Board of Governors uses Nilson data as one source in constructing the store card and gas card components of G.19 outstandings.

[9]See Bucks *et al*. (2006) for more details on the SCF.

[10]The modern SCF started in 1983 and began its triennial repeated cross-section sampling in 1989; as such, the only candidate for a pre-1989 comparison would be 1983. But the 1983 G.19 omits general purpose cards issued by financial institutions other than commercial banks, many important issuers of store cards, and charge cards (Nilson #339).

[11]The 1989–2004 SCFs use multiple imputation and provide five implicates for each household. I adjust point estimates for this by dividing each weighted-up estimate by 5, and correct standard errors per the standard procedure detailed in SCF codebooks.

3. Comparing Household and Industry Data

3.1. *Overview of Approach*

SCF measurement of credit card items—credit card account holding, charging, and borrowing—differs in potentially important ways from industry sources.[12] For each item the SCF plausibly produces narrower coverage, by design, than the G.19 or Nilson. The SCFs and industry measures may also be collected at different times of year. This section estimates how much discrepancy remains between SCF and industry sources after adjusting for differences in definition or coverage ("definition" below) and seasonality or other timing ("timing" below).

For each credit card item and each SCF survey year 1989–2004 I start by comparing the unadjusted discrepancy between the SCF and the main industry source for that item. This raw estimate does *not* adjust for definition or timing differences (this is the "casual" approach mentioned at the beginning of the paper): it simply takes the number featured in the G.19 release (e.g. the revolving debt "total") or a Nilson report (e.g. the total number of credit card accounts) and subtracts the weighted-up SCF aggregate for the related item. The ratio of the unadjusted industry number to the SCF number then gives the unadjusted "SCF Undercounting Factor."

I then adjust for definition and timing differences and recalculate the discrepancy and SCF Undercounting Factor. Tables report estimates for *combinations* of adjustments to save space.[13] I account for imprecision in the aggregates and the adjustments thereto using ranges of point estimates (minimal, moderate, and maximal) because the industry sources do not report standard errors, per standard practice with macro financial accounts.[14] The relevant SCF standard errors are small and I report some in the text to give the reader a sense of the SCF's precision.

3.2. *Accounts*

Table 2, Column 1 reports estimates of the unadjusted discrepancy between the total number of accounts reported in Nilson and calculated from the SCF, and the resulting SCF Undercounting Factor.[15] The unadjusted Undercounting Factor here ranges from a low of 1.6 in 1989 to a high of 2.4 in 2001 and 2004. The SCF estimates are precise; for example, the 2004 account total has an imputation-corrected standard error of 3 million (which is about 1/64 of the SCF point estimate). Columns 2–4 of Table 2 then present estimates of the discrepancy and Undercounting Factor that adjust for three potentially important definitional differences between the G.19 and SCF.

[12]Comparing credit limits and pricing might also be of interest, but I could not find industry data on credit limits, and pricing comparisons are greatly complicated by data limitations on both sides.

[13]For more details on each individual adjustment, please see the technical working paper version at http://www.dartmouth.edu/~jzinman/Papers/Zinman_MissingCardDebt_sep07.pdf.

[14]As Antoniewicz *et al*. (2005) note, "the complex structure of the macro financial accounts and the vast disparate sources that are used as inputs make calculating even the most simplistic standard error a daunting task."

[15]I report totals for general purpose (a.k.a. "bank") credit cards only to conserve space and because they generate the lion's share of charges and debt (Table 3).

TABLE 2

Missing Credit Card Accounts? Estimates of Discrepancy between Nilson and SCF

Year	Unadjusted Discrepancy (1)	Discrepancy After: Minimal Adjustments (2)	Moderate Adjustments (3)	Maximal Adjustments (4)
1989	59	53	26	−2
	1.6	*1.5*	*1.3*	*1.0*
1992	80	70	31	−8
	1.7	*1.6*	*1.3*	*0.9*
1995	162	146	71	−5
	1.9	*1.8*	*1.4*	*1.0*
1998	222	200	96	−9
	2.3	*2.1*	*1.5*	*0.9*
2001	273	237	116	−6
	2.4	*2.2*	*1.6*	*1.0*
2004	300	249	97	−56
	2.4	*2.2*	*1.5*	*0.7*

Notes: Discrepancies in millions of accounts.

Italics: SCF Undercounting Factor = (Nilson estimate/SCF estimate).

Columns 2–4 adjust for SCF top-coding, inactive accounts, and business-related accounts. See Section 3.2 of text for details and the technical working paper online for more details.

TABLE 3

Trends in Credit Card Holding and Usage

Year	(1) Proportion With Any Credit Card	(2) Proportion With Any General Purpose Credit Card	(3) Proportion of Charges on General Purpose Cards	(4) Proportion of Revolving on General Purpose Cards
1989	0.70	0.57	0.59	0.71
1992	0.72	0.63	0.64	0.74
1995	0.74	0.67	0.67	0.84
1998	0.73	0.68	0.70	0.87
2001	0.76	0.73	0.71	0.87
2004	0.75	0.72	0.80	0.89

Notes: All estimates from the SCF and weighted using variable x42001.

The first adjustment is for SCF top-coding. In each of the 1995–2004 surveys about 1 percent of the weighted sample has their number of accounts censored at 10. My minimal adjustment (Column 2) assumes that the average top-coded household has one additional account on average; the moderate and maximal adjustments (Columns 3 and 4) assume 5 and 10 additional accounts. This adjustment adds between 1 and 12 million accounts to the unadjusted SCF total.

The second adjustment is for inactive accounts. The unadjusted Nilson total includes them; SCF households may not report inactive accounts they have forgotten, closed,[16] or perceive as closed. Nilson reports totals for inactive accounts separately and I use this to make an adjustment. The minimal adjustment

[16]Issuers report an indeterminate number of closed accounts as active due to clerical errors.

assumes that households report all inactive accounts (and hence does not subtract anything from the unadjusted Nilson total), the moderate adjustment subtracts 50 percent of inactive accounts from the unadjusted Nilson total, and the maximal adjustment subtracts all inactive accounts from the unadjusted Nilson total (i.e. it assumes that SCF households do not report any inactive accounts). The effect of the adjustment for inactive accounts can be large, and is increasing over time; for example, Nilson reported 47 million inactive accounts in 1989 (29 percent of total accounts), and 230 million in 2004 (45 percent of total accounts).

The third adjustment is for business-related accounts. While the SCF prompts respondents to "not count . . . any business or company accounts," the unadjusted Nilson total includes them. Accordingly my minimal adjustment here subtracts an estimate of the count of open commercial accounts (from Nilson) from the unadjusted discrepancy; this adjustment ranges from 6 million in 1989 to 50 million in 2004. The moderate adjustment also subtracts a conservative estimate of personal accounts used for business purposes based on Nilson; this ranges from 4 million in 1989 to 32 million in 2004. The maximal adjustment uses a more aggressive estimate of personal accounts used for business purposes that ranges from 8 million in 1989 to 64 million in 2004.[17]

The net effect on the estimated SCF Undercounting Factor for credit card accounts varies across the three different sets of adjustments. The factor after minimal adjustments rises from 1.5 in 1989 to 2.2 in 2004, and is quite close in magnitude to the unadjusted Undercounting Factor in each year. The factor is much smaller and more stable after the moderate or maximal adjustments. The moderate-adjusted Undercounting Factor ranges from 1.3 to 1.6. The maximal-adjusted Undercounting Factor is 0.9 or 1.0 in every year except 2004, when it drops to 0.7.

3.3. *Recent Charge Volume*

Table 4, Column 1 reports estimates of the unadjusted discrepancy between monthly total credit card charge volume from Nilson (annual total/12) and the SCF (weighted-up charges on last bills),[18] and the resulting SCF Undercounting Factor. The unadjusted Undercounting Factor here is smaller than for accounts (or for debt, as detailed below). It ranges from 1.2 (in 1989) to 1.8 (in 1998), and was 1.5 in 2004. As with the other items the aggregate SCF estimates are precise; for example, 2004 charges have an imputation-corrected standard error of about $3.5 billion (which is about 1/30 of the SCF point estimate).

Columns 2–4 of Table 4 then present estimates of the discrepancy and Undercounting Factor that adjust for three potentially important definition and timing differences between the G.19 and SCF.

[17]The uncertainty about the correct point estimate for personal use of Nilson business-related accounts come in part from the fact that Nilson reports the number of business using personal cards, rather than the number of accounts.

[18]I report charge volume across all credit cards for comparability to the revolving debt estimates in the next sub-section (the G.19 does not disaggregate debt by card type, making it difficult to focus on general purpose cards only).

TABLE 4

Missing Credit Card Charges? Estimates of Discrepancy between Nilson and SCF

Year	Unadjusted Discrepancy (1)	Discrepancy After: Minimal Adjustments (2)	Moderate Adjustments (3)	Maximal Adjustments (4)
1989	5	2	–1	–4
	1.2	*1.1*	*1.0*	*0.9*
1992	13	12	6	–1
	1.4	*1.4*	*1.2*	*1.0*
1995	25	18	9	–1
	1.5	*1.4*	*1.2*	*1.0*
1998	41	28	14	0
	1.8	*1.5*	*1.2*	*1.0*
2001	49	30	12	–7
	1.6	*1.4*	*1.1*	*0.9*
2004	54	29	7	–15
	1.5	*1.3*	*1.1*	*0.9*

Notes: Discrepancies in nominal billions of dollars.

Italics: SCF Undercounting Factor = (Nilson estimate/SCF estimate).

Columns 2–4 adjust for seasonality, business-related charges, and cash advances. See Section 3.3 of text for details and the technical working paper online for more details.

The first adjustment is for seasonality.[19] While the Nilson estimate is based on annual total charges, SCF surveys are conducted May–December. Given that the SCF asks about charges on previous bills, this means that the SCF misses most of the Christmas driven uptick in charges. Although I could not find month-to-month charge data, Nilson shows that charges are typically about 10 percent higher in the fourth quarter (October–December) than in the average of the first three quarters. Thus if SCF data missed the fourth quarter uptick completely it would underestimate average monthly charges by about 3 percent. So I inflate the SCF number by 3 percent in my moderate adjustment, and by 1 percent and 5 percent in the minimal and maximal adjustments.

The second adjustment is for business-related charges. While the SCF prompts respondents to exclude business or company accounts, the Nilson year-end total includes charges on commercial accounts (which the SCF clearly means to exclude) and business-related charges on consumer cards (which the SCF may only partially exclude, if some accounts are used for both personal and business purposes). I use supplemental data from Nilson (on commercial account charges or business-related charges across all cards) and the Survey of Small Business Finances (on business-related charges on consumer cards) to estimate the appropriate adjustment. The maximal adjustment subtracts an estimate of all business-related charges from the Nilson unadjusted total. This is a substantial amount in all years; e.g. \$5 billion monthly = 14 percent of the Nilson total in 1989, and \$42 billion = 27 percent of the Nilson total in 2004. The minimal adjustment subtracts an estimate of charge volume for commercial accounts only. This too is substantial (at least relative to the unadjusted discrepancy) in all years; e.g. \$2

[19]Each credit card billing cycle (the time period covered by the SCF question) covers a month, so seasonality is the only *timing* adjustment is needed to make Nilson and SCF charges comparable.

billion monthly in 1989, and $24 billion in 2004. The moderate adjustment uses the midpoint of minimal and maximal adjustments for business-related charges.

The third adjustment is for cash advances. The Nilson total includes them; the SCF questions may exclude them if respondents do not think of cash advances as "charges." The maximal adjustment assumes that SCF respondents do not report any cash advances and hence subtracts an estimate of monthly cash advances from the unadjusted Nilson charge total. The cash advance estimate also comes from Nilson and is, for example, $2 billion in 1989 and $23 billion in 2004. The moderate and minimal adjustments subtract half and none of Nilson's cash advance volume.

On net the adjustments eliminate most or all of the discrepancy between charges reported in the SCF and Nilson. The estimated yearly SCF Undercounting Factor ranges from 1.1 to 1.5 with the minimal adjustments, from 1.0 to 1.2 with moderate adjustments, and from 0.9 to 1.0 with maximal adjustments. Charges seem to match up well under a range of assumptions.

3.4. *Borrowing*

Table 5, Column 1 reports estimates of the unadjusted discrepancy between the revolving "total" in the G.19 September 30 release and the related SCF item (balances remaining after last payment), and the resulting SCF Undercounting Factor. The unadjusted Undercounting Factor ranges from 2.9 (in 1989) to 3.8 (in 2001), and was 3.1 in 2004. As with the other items the SCF estimates are precise; e.g. 2004 total revolving balances have an imputation corrected standard error of $8 billion (which is about 1/30 of the SCF point estimate, and 1/67 of the

TABLE 5

Missing Credit Card Debt? Estimates of Discrepancy between G.19 and SCF

		Discrepancy After:		
Year	Unadjusted Discrepancy (1)	Minimal Adjustments (2)	Moderate Adjustments (3)	Maximal Adjustments (4)
1989	$132	$125	$67	$21
	2.9	*2.8*	*2.0*	*1.3*
1992	$179	$170	$103	$47
	2.9	*2.8*	*2.1*	*1.5*
1995	$297	$296	$195	$99
	3.3	*3.3*	*2.5*	*1.8*
1998	$390	$376	$269	$173
	3.2	*3.1*	*2.5*	*2.0*
2001	$522	$495	$372	$263
	3.8	*3.7*	*3.0*	*2.4*
2004	$537	$503	$362	$218
	3.1	*3.0*	*2.4*	*1.9*

Notes: Discrepancies in nominal billions of dollars.

Italics: SCF Undercounting Factor = (G.19 estimate/SCF estimate)

Columns 2–4 adjust for transaction balances, seasonality, non-credit card lines of credit, and business-related revolving on personal cards in the G.19. See Section 3.4 of text for details and the technical working paper online for more details.

unadjusted estimated discrepancy between the G.19 and SCF). Columns 2–4 then present estimates of the discrepancy and Undercounting Factor that adjust for four potentially important definition and timing differences between the G.19 and SCF.

A first adjustment to the G.19 revolving total is for timing. Most SCF surveys are conducted during May through December, but individual survey dates are not reported and hence one can not seasonally weight the reported credit card items. Hence I use different dates for the G.19 to help establish a range for comparison: December 30 (Column 2), September 30 (Column 3), and June 30 (Column 4). The upward trend in G.19 outstandings throughout the year (December 30 outstandings are on average 6 percent higher than September 30, and 8 percent higher than June 30) implies that the June 30 number typically produces the smallest discrepancy from the SCF, holding constant the other three adjustments detailed below.

The second adjustment removes recent charges from the G.19. Whereas the SCF revolving balance estimate includes only balances after the last bills were paid, the G.19 takes a snapshot of current debt outstanding. As such the G.19 includes not only the "revolving" balances captured by the SCF borrowing questions, but also "transaction" balances (basically, charges incurred by households since their previous billing cycles closed). To adjust for this, one first needs estimates of monthly charges. I construct these by simply adding an adjusted discrepancy estimated in Table 4 to unadjusted SCF monthly charges. Since the minimal (maximal) adjustment in Table 4 produces a larger (smaller) base of adjusted monthly charges, I use that minimal (maximal) adjustment to construct the maximal (minimal) adjustment for transaction balances needed in Table 5 for revolving debt. Next we need an estimate of the proportion of transaction balances that will show up in the G.19 snapshot of outstandings. The technical working paper shows how to estimate sensible upper and lower bounds for this proportion.[20] I use 1.35 for the upper bound: multiplying this by the maximal estimate of transaction balances, and then subtracting the product from the unadjusted G.19 revolving total, gives a maximal adjustment. This is adjustment is quite large (e.g. $48 billion and 36 percent of the unadjusted discrepancy in 1989; $198 billion and 37 percent in 2004). I use 1.0 for the moderate adjustment, and 0.5 for the lower bound and the minimal adjustment. Even the minimal adjustment reduces the unadjusted discrepancy substantially (e.g. by 10 percent in 1989 and 8 percent in 2004).

A third adjustment removes amounts from the G.19 revolving total that are not issued through credit cards but rather through prearranged overdraft plans or check-accessed lines of credit. (The SCF questions refer specifically and only to credit cards—there are separate questions for other lines of credit.) Non-credit card lines are *not* disaggregated in the G.19 releases but can be removed for 2001 and 2004, following Furletti and Ody (2006), by referencing the G.19's source

[20]Identifying a more precise estimate of the impact of transaction balances on G.19 outstandings would require data on the within-month distribution of charges and the duration of float obtained by convenience users (i.e. by those who do not revolve debt).

data.[21] The resulting point estimates for these non credit card lines are $32 billion in 2001 and $33 billion in 2004 (i.e. 6 percent of the unadjusted discrepancy).[22] But anecdotal evidence suggests that many of these "non-credit card" lines can in fact be accessed by card (as well as by check), so my minimal adjustment in Column 2 assumes that half of non-card lines can be accessed by cards and hence subtracts only half of the point estimate from the G.19 total. Column 3 subtracts the non-card point estimate. Column 4 allows for the possibility that my point estimates understate the amount of non-card outstandings—this is particularly important for earlier years, where my estimates are noisier and non-card lines may have been more prevalent—and subtracts 1.5 times the non-card point estimate from the G.19 total.

A fourth and final adjustment attempts to remove business-related revolving on personal cards from the G.19. Recall that while the G.19 includes *all* personal credit cards; in contrast the SCF instructs respondents to "not count . . . any business or company accounts." This implies that the G.19 includes some outstandings on personal cards that are used for business purposes and excluded by design from the SCF. I construct the minimal adjustment for such outstandings (in Column 2) using questions from the 1998 and 2003 Surveys of Small Business Finances (SSBFs) that ask: "On average, what is the balance of business charges on all owners' personal credit cards after payments are made?" I then linearly interpolate or extrapolate to get estimates for the other years in my sample. The SSBF numbers are quite small—only $1–7 billion, or 1–3 percent of unadjusted SCF revolving balances. These magnitudes are too conservative because the SSBF does not represent many types of businesses where personal cards are used. Nilson (various issues) finds that many large businesses provide cards that are in employees' names; these may be counted as personal cards in the G.19. Moreover the SSBF only *represents* the 6.3 million small businesses in Dun's Market Identifier file, while Nilson (#772) reports that over 20 million small business owners *used* personal cards for business purposes in 2002. Nilson (#776) also reports a total of $51 billion in business-related outstandings across both personal and commercial cards for year-end 2001. Since commercial cards are excluded from both the SCF and G.19 I use $51 billion to construct the maximal adjustment (Column 4), linearly interpolating or extrapolating with a triennial growth rate that matches the unadjusted G.19 revolving number. The moderate adjustment then uses the mean of the minimal and maximal.

The net effect of the adjustments on the magnitude and time pattern of the discrepancy between the G.19 and SCF varies across the three different sets of adjustments. The minimal adjustments range from 2.8 to 3.7 and produce discrepancies that are just slightly lower than the unadjusted discrepancy in each year. The moderate-adjusted discrepancies range from 2.0 to 3.0 and are

[21]Commercial banks report non-credit card lines separately in Call Report Schedule RC-C under RCONB539. Prior to 2001 this disaggregation did not exist in the Call Reports. Finance companies do not ever report non-credit card lines separately during my sample period, so I assume that they appear in the same proportion as for banks. Non-credit card unsecured revolving loans issued by thrifts or credit unions are *not* included in the revolving portion of the G.19 and therefore do not need to be subtracted here to produce comparable estimates. See Furletti and Ody (pp. 23–34 and appendix I) for more details.

[22]I linearly interpolate to get estimates for the years before 2001.

somewhat higher on average in the later years. The maximal-adjusted discrepancies range from 1.3 to 2.4 and are again somewhat higher on average in the later years.

4. Discussion: Sources of the Remaining Discrepancy in Credit Card Borrowing

Why do hundreds of billions of dollars in credit card debt remain missing, even after adjusting for definitional differences between the G.19 and SCF? Answering this question is critical for assessing the usefulness of household survey data on unsecured borrowing, and for designing and administering surveys that elicit useful data.

One possible source of the discrepancy is that researchers have overlooked something that induces issuers to overreport credit card assets in the G.19's source data. But as it stands the weight of the evidence points more toward SCF undercounting than G.19 overcounting.

Sample frame and *survey* response rate issues are not likely explanations for SCF undercounting, since as noted above the SCF seems to match up well with industry data on other balance sheet items and credit card charges, and relatively well on credit card credit accounts. But it remains possible that sample selection issues affect credit card balances more than other balance sheet items, and/or that undisclosed properties of the Nilson data make the relatively close alignment between SCF and industry measures of charges and accounts a mirage.

Nor does nonresponse on the SCF credit card borrowing *questions* seem to be a likely candidate. SCF staff report that they have scrutinized the credit card module closely and found no evidence of unusual (non)response patterns. This is not surprising given that the credit card questions come relatively early in the survey, minimizing the likelihood that survey fatigue leads respondents to supply easy answers like "zero."

Intentional underreporting of behavior that is viewed as socially undesirable is always a concern for household surveys (Wyner, 1980; Means *et al*., 1992). But the SCF is skilled at eliciting sensitive financial information, and evidence from Michigan Surveys of Consumers in 2001 (reported in Durkin, 2002) casts doubt on the hypothesis that credit card borrowing is widely stigmatized. Eighty-nine percent of credit card users reported that credit cards make managing finances "easier" or "no different," while only 10 percent reported "more difficult." Ninety percent of general purpose cardholders reported being "very satisfied" or "somewhat satisfied" overall with their card(s). Moreover it seems likely that if stigma played a role it would be declining over time as credit card use becomes more prevalent. Yet the Undercounting Factor for revolving balances in the SCF is higher in 2004 than 1989 by any of the four estimates in Table 5. So it seems likely that some type of *unintentional* underreporting plays an important role.

The impact of the sheer complexity of accounting for credit card borrowing throughout the household—most households have multiple cards, potentially used by several different household members—on reporting is important to consider. But complexity is ultimately unconvincing as an explanation for *under*reporting. Accounting for credit card charges requires *more* arithmetic than accounting for

current balances, yet SCF charges seem to match up well to the industry data. Other complex balance sheet items (e.g. on the asset side) also match up well with industry-side data. And even if respondents make unintentional reporting *errors* due to the private nature of credit card borrowing by other household members, it is hard to imagine why these errors would be biased in the direction of underreporting. Why not mean-zero, or *over*reporting errors?

A more novel explanation focuses on why respondents might systematically underestimate their *own* credit card balances. Results from a laboratory experiment suggest that consumers underestimate or forget credit card purchases because the act of paying by credit card is less memorable and painful than paying by check (Soman, 2001). Credit cards may stimulate spending as a result (see also Thaler, 1999), raising the intriguing possibility that a psychological bias affects not only reported but *actual* credit card borrowing behavior. But this explanation might struggle to account for the relatively complete SCF reporting of credit card *charges*.

In all there does not seem to be any obvious (or at least any single) explanation for SCF undercounting of credit card borrowing.

5. Is SCF Undercounting Worrisome?

As noted at the outset, SCF undercounting of credit card borrowing will be problematic for researchers if it is correlated with unobserved heterogeneity in preferences, resources, or another factor that is in turn correlated with outcomes of interest like financial condition, consumption paths, or portfolio choice. The variation over time in undercounting found in Table 5, Columns (3) and (4) is worrisome, particularly given the changes in credit card use and users documented in Tables 1 and 3. Table 3, Column 2 shows that many households (about 26 million) entered the credit card market from 1989 to 2004. The proportion of charges and revolving debt on general purpose cards rose sharply as well (Columns 3 and 4). Table 1 suggests that marginal cardholders (i.e. those with only one general purpose credit card, and hence those most likely to have entered the market and changed aggregate reporting behavior) were more likely to be female, nonwhite, unmarried, and have below median income in 2004 than in 1989. Debt reporting behavior may vary with demographic characteristics such as these; Karlan and Zinman (2008) find evidence suggesting that females (and perhaps relatively low-income borrowers) are less likely to report their expensive, unsecured debt.

Unobserved heterogeneity in respondent reporting of credit card debt could present a serious problem for researchers across a range of literatures and methods. It would confound attempts to estimate the distribution, incidence, and determinants of borrowing costs; this concern applies, for example, to the literature exploring why households simultaneously borrow expensively on cards while lending in low-yielding demand deposits. Studies in this vein using the SCF include Bertaut and Haliassos (2009), Gross and Souleles (2002), Telyukova (2008), Telyukova and Wright (2008), and Zinman (2007). An example of the inference problem in this literature is its attempt to answer the question of whether less-educated consumers are more likely to borrow high and lend low. If education is correlated with reporting behavior it will be unclear whether costs

are attributable to education *per se*, even in an otherwise well-identified model. Unobserved heterogeneity in reporting might also bias estimates of parameters or model fit; these concerns apply, for example, to literatures estimating the effect of debt burden on financial distress. Calem *et al*. (2006) use the SCF to estimate the effect of credit card borrowing on financial distress, conditional on a rich set of control variables. The concern here is that underreporting of credit card borrowing that is correlated with incompletely observed drivers of distress (e.g. risk aversion, impatience, access to other sources of liquidity) will bias the estimate of borrowing's causal effect in an indeterminate direction.[23] It might also affect studies of the effect of house prices on consumption,[24] and the fit of alternative models of intertemporal and portfolio choice, e.g. computational life-cycle models increasingly use credit card borrowing as a key outcome to be fit and focus on particular sub-samples (cohorts, education groups, financial market participants, etc).[25]

In all, the available evidence is quite far from permitting definitive inference on whether SCF data on credit card borrowing is useful for researchers. But the pattern of evidence is consistent with changes in the SCF Undercounting Factor being driven by changes in credit card use and/or users that could well be correlated with confounding unobserved heterogeneity. This disconcerting possibility highlights the importance of identifying the source(s) of the remaining discrepancy between SCF and industry measures of credit card debt. The next section discusses possibilities for future work along these lines.

6. Conclusion

Many government agencies, international organizations, and researchers (including the author) conduct household surveys that include credit cards or other high-frequency borrowing. As interest in household finance continues to grow, demand for survey data on these important sources of credit is likely to increase.

I create comparable measures of aggregate credit card use based on household data from the Survey of Consumer Finances (SCF), and industry data from the G.19 and Nilson Reports. The SCF and industry sources match up well on credit card charges, fairly well on account totals, and not especially well on revolving debt. My estimated lower bound for 2004 debt discrepancy is $200 billion. The available evidence is far from definitive but suggests that the SCF misses half of revolving credit card debt.

Perhaps more worrisome is that I also find some evidence of growth over time in the debt discrepancy. This growth parallels substantial changes in credit card use (the shift to general purpose cards) and credit card users (more marginal credits

[23]See Karlan and Zinman (2008) for related evidence from a different setting. Many studies review the growing interest in the determinants and effects of household debt burden; see, e.g. Debelle (2004), Dynan and Kohn (2007), and White (2007).

[24]Here unbiased measures of liquidity constraints (and hence presumably of available credit under credit card limits) are needed to identify or interpret parameters of interest; see, e.g. Campbell and Cocco (2007) and Hurst and Stafford (2004).

[25]Unobserved heterogeneity in reporting behavior across these sub-samples could produce misleading estimates of model fit or calibrated/inferred parameters. Studies in this vein using the SCF include Angeletos *et al*. (2001), Carroll (2001), Davis *et al*. (2006), and Laibson *et al*. (2009).

entering the pool). The parallel suggests that any SCF undercounting may be correlated with unobserved heterogeneity in household characteristics that could confound inference on the relationship between credit card borrowing and outcomes of interest like household financial condition, consumption paths, and portfolio choice. If the SCF—a gold standard for eliciting comprehensive and accurate financial information—is problematic, then other household surveys likely face even greater problems. Given this disconcerting possibility it is critical to develop sharper evidence on whether and why household surveys actually undercount credit card borrowing.

One approach is to continue scrutinizing the G.19 and other industry sources. For example, part of the discrepancy could be due to changes in the composition of credit card debt issuers and hence in their data reporting (see, e.g. Furletti and Ody, 2006, appendix F). Another approach is to collect richer data within existing reporting instruments. The reconciliation exercises in this paper highlight several additional pieces of survey and industry data—e.g. on credit cards for business use, on months with unusually high charges, on the within-month time path of charges—that could be used to obtain more precise estimates of the magnitude and source(s) of the discrepancy between household and industry measures of credit card borrowing.

Future work would also do well to focus on validation studies using lender data and/or incentive compatible elicitation procedures (e.g. a game that rewards subjects for accurate estimation of a credit card balances, where the subject herself is responsible for documenting the actual balance). These methods would permit sharper tests of whether consumers actually underreport systematically. Coupled with supplemental data collection on decision inputs (e.g. preferences, expectations, problem-solving ability, liquidity constraints) they would also shed light on what drives any underreporting, and on whether these drivers are likely to confound inference using credit card borrowing data from household surveys.

References

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg, "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, 15(3), 47–68, 2001.

Antoniewicz, Rochelle, Riccardo Bonci, Andrea Generale, Grazia Marchsese, Andrea Neri, Karen Maser, and Patrick O'Hagan, "Household Wealth: Comparing Micro and Macro Data in Canada, Italy, and United States," Working Paper, Board of Governors of the Federal Reserve, January 18, 2005.

Bertaut, Carol, Michael Haliassos, and Michael Reiter, "Credit Card Debt Puzzles and Debt Revolvers for Self Control," *Review of Finance*, 2009, forthcoming.

Brackstone, Gordon, "Statistical Uses of Administrative Data: Issues and Challenges," *Proceedings of Statistics Canada Symposium*, 5–16, 1987.

Bucks, Brian, Arthur Kennickell, and Kevin Moore, "Recent Changes in Family Finances: Evidence from the 2001 and 2004 Survey of Consumer Finances," *Federal Reserve Bulletin*, 92, A1–38, February 2006.

Calem, Paul, Michael Gordy, and Loretta Mester, "Switching Costs and Adverse Selection in the Market for Credit Cards: New Evidence," *Journal of Banking and Finance*, 30, 1653–85, 2006.

Campbell, John Y. and Joao Cocco, "How Do House Prices Affect Consumption? Evidence from Micro Data," *Journal of Monetary Economics*, 54, 591–621, 2007.

Carroll, Christopher D., "A Theory of the Consumption Function, With and Without Liquidity Constraints," *Journal of Economic Perspectives*, 15(3), 23–45, 2001.

Davis, Steven, Felix Kubler, and Paul Willen, "Borrowing Costs and the Demand for Equity Across the Life-Cycle," *Review of Economics and Statistics*, 88, 348–62, 2006.
Debelle, Guy, "Household Debt and the Macroeconomy," *BIS Quarterly Review*, March 2004.
Durkin, Thomas, "Consumers and Credit Disclosures: Credit Cards and Credit Insurance," *Federal Reserve Bulletin*, 88, 201–13, 2002.
Dynan, Karen and Donald Kohn, "The Rise in U.S. Household Indebtedness: Causes and Consequences," FEDS Working Paper, 2007.
Edelberg, Wendy, "Risk-Based Pricing of Interest Rates for Consumer Loans," *Journal of Monetary Economics*, 53, 2283–98, 2006.
Elliehausen, Gregory and Edward Lawrence, "Payday Advance Credit in America: An Analysis of Customer Demand," Credit Research Center, Georgetown University, April 2001.
Furletti, Mark and Christopher Ody, "Measuring U.S. Credit Card Borrowing: An Analysis of the G.19's Estimate of Consumer Revolving Credit," Federal Reserve Bank of Philadelphia Payment Cards Center Discussion Paper, April 2006.
Gross, David and Nicholas Souleles, "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data," *Quarterly Journal of Economics*, 117, 149–85, 2002.
Hurst, Erik and Frank Stafford, "Home Is Where the Equity Is: Mortgage Refinancing and Household Consumption," *Journal of Money, Credit, and Banking*, 36, 985–1014, 2004.
Johnson, Kathleen, "The Transactions Demand for Credit Cards," *The B.E. Journal of Economic Analysis & Policy: Topics*, 7(1), Article 16, 2007.
Karlan, Dean and Jonathan Zinman, "Lying About Borrowing," *Journal of the European Economic Association Papers and Proceedings*, 6(2–3), August 2008.
Laibson, David, Andrea Repetto, and Jeremy Tobacman, "Estimating Discount Functions with Consumption Choices Over the Lifecycle," *American Economic Review*, 2009, forthcoming.
Means, Barbara, Katherine Habina, Gary Swan, and Lisa Jack, "Cognitive Research on Response Error in Survey Questions on Smoking," *Vital Health & Statistics*, Series 6, Number 5, National Center for Health Statistics, Hyattsville, MD, 1992.
Soman, Dilip, "Effects of Payment Mechanism on Spending Behavior: The Role of Rehearsal and Immediacy of Payments," *Journal of Consumer Research*, 27, 460–74, 2001.
Sullivan, James X., "Borrowing During Unemployment: Unsecured Debt as a Safety Net," *Journal of Human Resources*, 43, 383–412, 2008.
Telyukova, Irina, "Household Need for Liquidity and the Credit Card Debt Puzzle," Working Paper, University of California–San Diego, December 2008.
Telyukova, Irina and Randall Wright, "A Model of Money and Credit, with Application to the Credit Card Debt Puzzle," *Review of Economic Studies*, 75, 629–47, 2008.
Thaler, Richard, "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12, 183–206, 1999.
White, Michelle, "Bankruptcy Reform and Credit Cards," *Journal of Economic Perspectives*, 21, 175–99, 2007.
Wyner, Gordon, "Response Errors in Self-Reported Number of Arrests," *Sociological Methods and Research*, 9, 161–77, 1980.
Zinman, Jonathan, "Household Borrowing High and Lending Low Under No-Arbitrage," Working Paper, Dartmouth College, April 2007.