NBER WORKING PAPER SERIES

ASSESSING NONRESPONSE BIAS IN MACRO INDICATORS BY COMBINING PARA-, ADMINISTRATIVE, AND SURVEY DATA

Ori Heffetz
Guy Lichtinger
Daniel B. Reeves

Working Paper 32935
http://www.nber.org/papers/w32935

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2024

We thank participants in the Cornell Behavioral Economics Reading Group (BERG), and Jerusalem Economics, Society, and Culture (JESC) lab. We thank Rebecca Krieger for responding to our questions regarding the Israel Central Bureau of Statistics (ICBS) procedures. Neta Bar Ziv, Matan Gibson, Ophir Halperin, Lev Maresca, Tal Asif, and Gabriela Cohen Hadid provided invaluable research assistance. We also thank Nikolas Mittag and Michael Khalfin for their helpful comments. This study was supported by the Maurice Falk Institute for Economic Research in Israel. It was performed in the ICBS research room using de-identified Microdata (direct identifiers removed from data) in files prepared specifically for this project. The authors have no financial or other material interests related to this research to disclose. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Assessing Nonresponse Bias in Macro Indicators by Combining Para-, Administrative, and Survey Data
Ori Heffetz, Guy Lichtinger, and Daniel B. Reeves
NBER Working Paper No. 32935
September 2024
JEL No. C18, C83, J60

ABSTRACT

National surveys are crucial for estimating key economic aggregates, including the unemployment rate, labor force participation, and household expenditures. The accuracy of these indicators is increasingly under scrutiny due to declining response rates and the consequent risk of nonresponse bias. How can we assess nonresponse bias in these key economic aggregates?

Using two Israeli national surveys, we propose a novel approach. First, using often-available paradata such as number of contact attempts and nonresponse reason, we create respondent and nonrespondent subcategories. Second, using rarely-available merged administrative records, we identify, for each nonrespondent subcategory, which respondent subcategory appears to resemble it most. We find that nonrespondents are a heterogeneous group: some—e.g., those temporarily unavailable—share administrative-record demographic and outcome profiles with harder-to-reach respondents, while others—e.g., refusals and withdrawals—are more similar in the administrative data to easier-to-reach respondents. Third, assuming that these resemblances would extend to survey outcomes, we impute (always-available) survey-based aggregates to nonrespondents within each paradata-based subcategory.

We demonstrate that our method can help assess nonresponse bias in surveys lacking matched administrative records, using the U.S. Consumer Expenditure Survey as an example.

Ori Heffetz
S.C. Johnson Graduate School of Management
Cornell University
324 Sage Hall
Ithaca, NY 14853
and The Hebrew University of Jerusalem
and also NBER
oh33@cornell.edu

Guy Lichtinger
Department of Economics
Harvard University
guylichtinger@g.harvard.edu

Daniel B. Reeves
Cornell University
Capital One
1600 Capital One Dr, 448F
McLean, VA 22102
dbr88@cornell.edu

Introduction

Despite the abundance of big data and automated data collection methods, survey data remains the primary and often the exclusive source for a wide range of national statistics, social indicators, and economic studies. Closely monitored by global markets and extensively utilized in policy and research, macroeconomic indicators like national unemployment rates are derived from labor-force surveys, and inflation figures are based on weights constructed from household expenditure surveys.

A worldwide trend of declining survey response rates over recent decades has therefore raised significant concern (Meyer et al. (2015); Williams and Brick (2018); Luiten et al. (2020); and Rothbaum and Bee (2021)). In the U.S., for example, the response rate for the widely used Current Population Survey (CPS), the primary source of many labor force statistics, has dramatically dropped from 90 to 70 percent within just a decade (2013–2023). The Consumer Expenditure Survey (CE), another major national survey that serves as the source for the Consumer Price Index's market basket, has experienced an even more drastic decline during this period, from nearly 70 to approximately 40 percent (U.S. Bureau of Labor Statistics (2023b)).

Even if these trends plateau, the current nonresponse levels—30 percent in the CPS and 60 percent in the CE—are alarming. When units (individuals or households) from the sampled population do not participate (i.e., they are "nonrespondents"), then unless these nonrespondents (in aggregate) are statistically identical to participants ("missing at random" or, at least, random conditioned on observables), survey-based outcomes could suffer from nonresponse bias. Have increasing nonresponse rates turned this theoretically well-studied possibility into an actual empirical problem? Have indicators on the state of national economies, labor markets, and households become less representative of the populations they are meant to track? In this paper, we offer a new method that can potentially help answer these questions.

The basic idea behind our proposed method is simple: Nonrespondents are not all alike. Importantly, paradata can tell us a lot about them. Think for example of a person classified as a "Language Difficulty" nonrespondent. People with language difficulty often experience labor-market difficulty. Thus, even without interviewing them, knowing someone's nonresponse reason could mean knowing something about, e.g., their labor-force situation, income, and total expenditures. In addition, administrative records—when available—can turn this knowledge useful for estimating nonresponse bias. Specifically, as we demonstrate, administrative records can help identify subgroups among respondents that appear similar to certain categories of nonrespondents, and use those respondents' survey data to assess the accuracy of economic aggregates.

Existing empirical research assessing the accuracy of survey data has mostly taken one of two approaches. The first approach involves leveraging administrative records, which cover both respondents and nonrespondents, and comparing them with survey results.[1] However, this approach is only applicable to surveys whose outcomes can be verified, at least to some degree, in administrative records. Naturally, many surveys do not satisfy this condition—or else, they would not be in such wide use in the first place. Furthermore, even when outcomes are in principle reasonably replicable using administrative records, access to such records is often limited, due to privacy concerns, legal barriers, and other reasons.

The second approach, employed by a smaller body of literature, focuses on surveys whose paradata include information on the difficulty of reaching each respondent (e.g., the number of visits or phone-call attempts made before completing the interview). By comparing outcomes of easy-to-reach respondents with those of harder-to-reach ones, researchers can make inferences about potential nonresponse bias, under the hypothesis that nonrespondents—who are, in principle, harder-to-reach than the hardest-to-reach

[1] In a recent example, Meyer and Mittag (2019) focus on responses in the Current Population Survey (CPS) regarding amounts received from four government income-transfer programs (e.g., food stamps and subsidized housing). Since most transfer programs record recipients' Social Security Number (SSN), the authors are able to link responses to administrative records, and compare reported to actual amounts received at the household level. They have a high match rate: they can identify the SSN of at least one household member in 91 percent of CPS households in their sample. Their findings suggest that using survey data sharply understates the income of poor households. In a subsequent paper, Meyer and Mittag (2021) use linked administrative data to demonstrate that survey errors in government transfer information are primarily due to measurement errors rather than coverage errors (which include survey frame limitations and nonresponse bias). Other studies using this approach include Kreuter et al. (2008); Bollinger et al. (2019); Davern et al. (2019); and Küfner et al. (2022).

respondents—extend, out of sample, difficulty-of-reaching trends that are found within sample.[2] Since this approach is based only on data from respondents, it could potentially assess certain types of nonresponse bias in any survey. However, this approach faces an inherent limitation of survey-based research: the lack of direct data on nonrespondents and, consequently, the inability to directly evaluate hypotheses about them.

In this paper, we employ the first approach to address the limitation of the second approach. Specifically, we use administrative records to replace the above hypothesis with empirical evidence regarding the similarities between respondent and nonrespondent subgroups. We analyze two prominent Israeli national surveys that closely resemble the CPS and CE in terms of outcomes and methodologies. The Israeli surveys have detailed paradata on both respondents and nonrespondents and, crucially, their data can be matched with administrative records, containing proxies for important survey outcomes. This matching allows us to study the nonrespondents directly and compare them with the respondents, both collectively and across the difficulty-to-reach spectrum.[3] Our objective is to look for systematic patterns that could empirically inform researchers' use of respondent and nonrespondent paradata, available in many surveys, to extrapolate information about nonrespondents from respondent data. If patterns that we find—and that replicate across the two surveys—hold more generally, then we may be able to make inferences about nonrespondents in other surveys where, unlike in our surveys, administrative records on nonrespondents are not available.

[2]In a recent example, Heffetz and Reeves (2019) focus on three major U.S. surveys (the CPS, BRFSS, and CE), examining four widely used indicators: unemployment rate, labor force participation rate, obesity rate, and household expenditures. They reveal robust differences between easy-to-reach and hard-to-reach respondents across all outcomes, even after adjusting for demographic differences across the groups. Thus, under the hypothesis that the nonrespondents continue the within-sample trends, their study suggests that nonrespondents may differ significantly from respondents in ways that standard weighting procedures do not address. Other studies using the difficulty-to-reach approach include Curtin et al. (2000); Heffetz and Rabin (2013); Behaghel et al. (2015); and Meterko et al. (2015).

[3]We are aware of two papers investigating difficulty trends where true outcomes are known for respondents and nonrespondents from administrative records. Lin and Schaeffer (1995) match child-support transfer data with a telephone survey and find that nonrespondents do not resemble hard-to-reach more than easy-to-reach respondents. In contrast, Kreuter et al. (2010) link administrative and survey data for German unemployment-benefits recipients and find that adding hard-to-reach respondents reduces nonresponse bias. However, both studies treat nonrespondents as a homogeneous group. As we show below, distinguishing between different nonrespondent subgroups is key to our findings and conclusions.

In Section 1, we introduce our survey and administrative records. We study two widely used national surveys, conducted by the Israel Central Bureau of Statistics (ICBS): the Labor Force Survey (LFS) and the Household Expenditure Survey (HES). They provide the official Israeli estimates for four main outcomes we focus on in this paper: labor force participation (LFS), the unemployment rate (LFS), household incomes (HES), and household expenditures (HES).[4]

For both LFS and HES, the ICBS successfully matched a significant majority of the sampled dwellings (93.8 and 89.6 percent, respectively), including both respondents and nonrespondents, with individual-level administrative records. This matching was facilitated by the national Dwellings and Buildings Registry, which links dwellings with their residents. The administrative records include detailed demographic information (age, sex, education), along with outcome-related data such as months of employment and income. Based on this data, we construct administrative proxy variables representing our surveys' outcomes of interest. In the LFS sample, we employ two administrative proxies for labor force participation status: an "any reported income" indicator, indicating a successful match with the Individual Income Registry (implying non-zero employment within the tax year), and a "full annual employment" indicator, indicating employment in all twelve months of the year. For the HES, our primary outcome of interest is administratively recorded individual-level income, as reported to the Treasury of Israel (we focus on income because expenditures are not recorded in the administrative records).

In Section 2, we analyze demographic and outcome patterns among respondents based on the difficulty of reaching them, using both survey and administrative data. The results from both data sources are consistent and align with past findings from the U.S. Hard-to-reach respondents tend to be younger and more educated. Moreover, before controlling for these demographic differences, we also find significant, U.S.-like trends in our outcomes of interest: hard-to-reach respondents are more likely to participate in the labor force, experience less unemployment, and report higher income and expenditures. Interestingly, after adjusting for demographic differences across the hard- and easy-to-

[4]These surveys are also used in economics research, e.g., Ben-Porath and Gronau (1985), Finkel et al. (2006), Schechtman et al. (2008), and Yashiv and Kasir (2013).

reach, these outcome differences become less pronounced than in the (adjusted) U.S. data; however, qualitatively, they remain mostly similar.

In Section 3, we utilize administrative records to examine nonrespondents and compare them with respondents. In Section 3.1, we use paradata to classify nonrespondents into subgroups based on their reasons for nonresponse, as recorded by the ICBS. This categorization reveals that the nonrespondents are a heterogeneous group. Some subgroups—those currently absent or faced with temporary difficulties (e.g., time constraints or temporary illness)—have similar demographics and outcomes to the hardest-to-reach respondents, suggesting they can indeed be viewed as respondents so difficult that they remained out of reach. In contrast, other subgroups—those who refused to participate, withdrew in the middle of the process, or were unable to participate due to permanent or language difficulties—appear different from the typical respondent in other ways than merely requiring more contact attempts to be reached. In fact, their demographics and outcomes are more similar to those of the *easiest-to-reach* respondents and appear even "easier" than them in some cases. Section 3.2 then categorizes nonrespondents by their number of (ultimately unsuccessful) contact attempts. We find clear trends among nonrespondents in both demographics and outcomes as the number of contact attempts increases, similar to the trends observed for respondents. This might indicate that interviewers make more contact attempts to nonrespondents they believe just need additional attempts to be reached, and fewer attempts to those differing from respondents in other ways. Section 3.3 examines the overall level of nonresponse bias in the survey, when measured using administrative outcomes, and finds limited bias.

In Section 4 we demonstrate how the insights from Section 3 can help assess nonresponse bias in the survey outcomes when administrative data is not available. This approach relies on two main (often-available) components: paradata on the distribution of nonrespondents by nonresponse category and survey outcomes of respondents by difficulty of reaching. Based on our comparisons of nonrespondents to respondents from Section 3.1, we extrapolate outcomes for the nonrespondent categories using the outcomes of respondents from the most similar difficulty category. In section 4.1, we present our nonresponse bias assessment for our four main survey outcomes. Our analysis indicates limited bias in these outcomes, which aligns with our findings using administrative data. Finally, in Section 4.2 we demonstrate how our method could potentially help assess nonresponse bias in surveys lacking matched administrative records, using the U.S. Consumer Expenditure Survey (CE) expenditure data as an example. Our results provide suggestive evidence of an upward nonresponse bias in the CE expenditure data, ranging from 1.7 to 8.4 percent, depending on the assumptions used.

In Section 5, we conclude and offer three practical insights. First and foremost, as we illustrate in Section 4, information on nonrespondents' nonresponse reason distribution, combined with data on respondents categorized by difficulty to reach, could help assess potential nonresponse bias in aggregate indicators. Second, such paradata on both respondents and nonrespondents should be routinely recorded, maintained, and, importantly, made readily accessible to survey users. Unlike the CE, many other national surveys do not currently publish detailed paradata that could facilitate assessing potential nonresponse bias using our methods (including, e.g., the CPS). Finally, our research contributes to an existing body of work suggesting that focusing solely on increasing response rates to reduce nonresponse bias may not achieve the intended results (e.g., Curtin et al. (2000); Keeter et al. (2000); Groves (2006); and Küfner et al. (2022)). Nonrespondents are a heterogeneous group with diverse demographics, outcomes, and nonresponse reasons. Different approaches, methods, and levels of effort to increase response rates are likely to differentially convert different subgroups of nonrespondents, thereby influencing the nature and extent of nonresponse bias. In particular, efforts to reduce nonresponse by merely increasing contact attempts may disproportionately convert temporarily unavailable nonrespondents, inadvertently moving the resulting profile of respondents *away* from that of the entire population that includes nonrespondents facing more permanent challenges.

1 Data

1.1 Survey Data and Paradata

1.1.1 Labor Force Survey (LFS)

The Labor Force Survey (LFS) is a monthly survey conducted by the Israel Central Bureau of Statistics (ICBS).[5] Similar to the U.S. Current Population Survey (CPS), it is the country's official source for the unemployment and labor force participation rates. In this paper we analyze data from the LFS collected from 2012 to 2017. We exclude pre-2012 data due to major changes to the survey's frequency and methodology, while post-2017 data was not available at the start of our study.

The LFS sample is selected once a year using a two-step sampling process. First, localities in Israel are sampled with probability proportional to their population. Then, using the Dwellings and Buildings Registry (described in more detail in section 1.2.3 below), dwellings are sampled from each selected locality.[6] The survey's sampling frame does not include households living outside government-recognized localities, such as those permanently living in army bases without a civilian address.

When dwellings are sampled, ICBS interviewers initiate contact. In cases where the first contact attempt fails, interviewers usually make at least two additional attempts during the two-week window after the first contact. Each unsuccessful attempt is documented along with its corresponding reason. If no successful contact was made within this timeframe, the household is classified either as a nonrespondent or as ineligible. Nonrespondents consist of dwellings that do not participate due to reasons such as absence, refusal, language issues, or difficulties in locating them. Ineligible dwellings primarily include dwellings that are nonresidential, unpopulated, used only occasionally, or those that do not belong to the intended survey population. The ICBS excludes these ineligible dwellings from the survey data and does not factor them into the calculations

[5] LFS information below is based on ICBS (2019).

[6] A small part of the population (e.g., Bedouin tribes and institutions such as nursing homes) is covered by a permanent sample taken from the 2008 Census.

for nonresponse rates. We therefore exclude them from our analysis.

Respondents are asked to participate in four consecutive months of interviews, followed by an eight-month break, and then four more months of interviews (the same 4-8-4 pattern is used by the U.S.'s CPS). The initial interview, known as the Stage A interview, is primarily conducted face-to-face. However, specific situations such as adverse weather, interviewer shortages, or a participant's request can lead to a phone interview. The subsequent seven interviews (Stages B–H) are typically phone interviews.

We focus on the Stage A interview because it is the initial contact with households, and it is typically conducted in person. In that relatively short interview (about 30 minutes), an adult household member provides information about all household members aged 15 and above. The survey data we received from the ICBS includes demographic and labor force information about these individuals. We are particularly interested in the labor force participation and unemployment questions, which align with the Organization for Economic Co-operation and Development (OECD) definitions. Household members are considered employed if they worked for at least one hour during the week preceding the interview or if they were temporarily absent from work during this week due to illness, vacation, reserve army service, etc. If they did not work during that week (and it was not defined as a temporary absence) and actively sought employment in the four weeks prior to the interview, they are classified as unemployed. The labor force includes individuals who are either employed or unemployed.

Beyond the survey data, the ICBS provided us with paradata regarding the Stage A interview. This paradata contains details about the attempts made to establish contact with each dwelling.[7] For dwellings that were not reached, the paradata includes the reasons for non-contact as determined by the interviewer.

[7]The paradata provided by the ICBS is at the dwelling, rather than the household level. However, almost all dwellings (more than 99 percent) contain only one household.

1.1.2 Household Expenditure Survey (HES)

The Household Expenditure Survey (HES) is conducted annually by ICBS to analyze the expenses, consumption, and income of Israeli households.[8] It is used to determine the weights for the Consumer Price Index's consumption basket, calculate the income distribution, and establish the poverty threshold. Our analysis is based on the 2012–2016 HES data. Similar to the LFS, we excluded pre-2012 data due to substantial changes in the survey's methodology, while post-2016 data was unavailable at the start of the study.

The HES sample population represents the general population of Israel, with the exception of households residing outside of recognized localities, much like the LFS. However, the HES additionally excludes collective kibbutzim and the Bedouin communities. The HES sampling process mirrors the LFS's: it begins by selecting localities, and then picks dwellings from those localities using the Dwellings and Buildings Registry.[9]

Similarly to the LFS, ICBS interviewers typically make several contact attempts with a dwelling before ceasing efforts. If they cannot establish contact, the dwelling is classified either as a nonrespondent or as ineligible. The contact window is significantly longer than the LFS's two-week window. Among HES dwellings with multiple contact attempts, 32 percent had a gap of more than 14 days between the initial and final attempt, and for 15 percent, this gap extended beyond two months.

Each participating household goes through three stages. In the first stage, an in-person interview is conducted to obtain basic demographic and economic data about each household member. The second stage involves a diary, where households record in detail their daily expenses over a two-week period (a similar method is used by the U.S.'s CE diary survey).[10] The final stage is a questionnaire that assesses significant or unusual expenditures and income based on reports covering either a 3-month or 12-month period

[8] HES information below is based on ICBS (2018).

[9] A very small portion of the HES sample is drawn from "complementary samples," including dwellings from locations not covered by the Dwellings and Buildings Registry, such as student dormitories and immigrant absorption centers.

[10] Participants who withdrew before or during the diary phase are categorized as nonrespondents (making up 31 percent of HES nonrespondents).

preceding the interview date (depending on the purchase frequency of the item).

Unlike the LFS, the HES data is at the household rather than the individual level. This data includes households' total expenditures and income (including income from labor and other sources, such as government transfers), as well as demographic information about the household head (defined as the member who works the most hours).

As with the LFS, the ICBS provided us with dwelling-level HES paradata. This paradata includes information about the number of visits until the initial successful contact or the final attempt, and the nonresponse reason for dwellings, where applicable.

1.2 Administrative Data

1.2.1 Administrative Sources

We rely on three administrative data sources, each updated annually by ICBS. First, the Individual Income Registry serves as our source for administrative records on income and labor force participation. It is based on tax reports submitted to the Israeli Treasury. The registry provides yearly income data for all working individuals, including both employees and the self-employed. For employees, the registry includes a variable indicating the number of months worked during the tax year, calculated based on reported salaries.

Second, the Education Registry is based on data from various sources. Our analysis focuses on the number of years of education for each individual. It is worth noting that even with accurate survey responses, the LFS and HES survey data may not always match the Income and Education registry data. For instance, individuals working in informal jobs or those who received education from informal institutions may report this information in the surveys, but it might not appear in the administrative records.

Lastly, the Resident Registry is a database that contains demographic information on all residents of Israel. It serves as our administrative source for variables such as age, sex, and ethnicity (Jewish vs. non-Jewish). The data is gathered from Israeli birth certificates and reports submitted to the government from those born abroad.

1.2.2 Administrative Outcome Variables

Based on the Individual Income Registry, we construct proxy variables for the surveys' outcomes of interest. To capture labor force participation status, we use two proxies: an "any reported income" proxy that indicates a match to the income registry (implying some level of employment within the tax year), and a "full annual employment" proxy, indicating employment in all twelve months of the year.[11] Note that the Individual Income Registry does not directly capture information on unemployment since non-employed individuals are only classified as unemployed if they are actively seeking employment, which is not readily recorded in administrative records. Nevertheless, the "any reported income" proxy plausibly proxies the labor force participation rate, as most unemployed individuals (who, by definition, were actively seeking employment) are highly likely to have been employed for at least part of the tax year.

For the HES, we rely on administratively documented individual-level income as the primary administrative outcome. Expenditures are not recorded in the administrative records; in our survey data, their correlation with income is 0.62.

1.2.3 Data Matching Process

To investigate nonresponse bias, administrative information is matched to most dwellings sampled for the LFS and HES, including both respondents and nonrespondents (described further in Section 1.3). This matching is performed by the ICBS using the Dwellings and Buildings Registry, which contains dwelling information collected from annual property-tax payment reports, including personal ID numbers of both the property-taxpayer and the dwelling owner.[12] We focus on matching data with the taxpayer, typically the resident, rather than the property owner, who may not reside in the dwelling, especially in rental situations. This implies that for each dwelling, the matched ad-

[11]A limitation of the second proxy is that it incorrectly classifies the self-employed as not working. This is true also for women in maternity leave, who are considered part of the labor force, but will not be captured by this proxy.

[12]The Israeli personal ID number is a unique 9-digit number issued by the Ministry of Interior Affairs to every Israeli. In some ways, it functions similarly to the Social Security Number in the U.S.

ministrative information pertains to the property-tax payer, rather than each household member (as in the LFS data), the household head (as in the HES demographic data), or the total household (as the HES income and expenditure data).

The matching process was based on the year in which the survey was conducted. For example, if the dwelling was sampled to be surveyed in 2014, the ID of the registered taxpayer of that dwelling in 2014 was taken and matched to all three registries using the data from 2014. One limitation of this process is that there could be a timing mismatch between the surveyed occupant and the registry information if the occupants changed between the time the registry information was collected and when the survey was conducted.

1.3 Sample Preparation

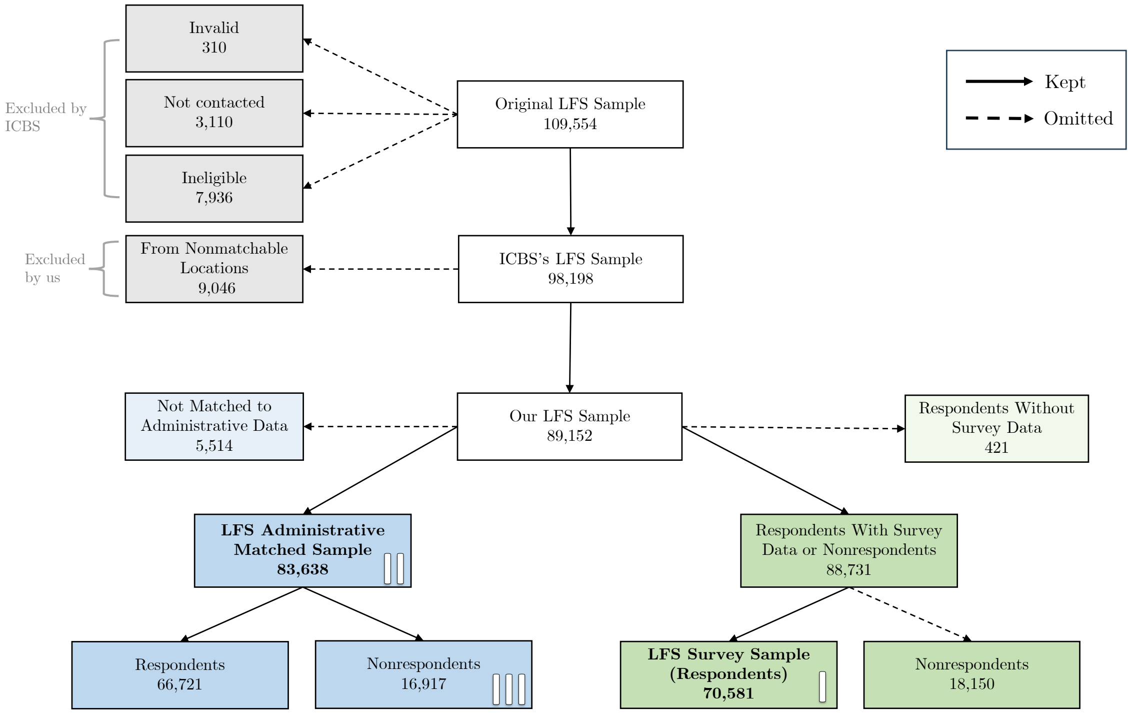
Figures 1 and 2 present our sample preparation process. To analyze patterns among respondents based on their survey data, we use the "LFS Survey Sample" (box I) and "HES Survey Sample" (box IV). To examine patterns among both respondents and non-respondents based on their administrative records, we employ the "LFS Administrative Matched Sample" (Box II) and the "HES Administrative Matched Sample" (Box V). The rest of this section provides a detailed explanation of these figures.

1.3.1 LFS Samples

Figure 1 outlines the creation of the LFS samples. The initial sample includes 109,554 dwellings, representing all dwellings sampled from 2012 to 2017. From this sample, we exclude 310 invalid dwellings (as flagged by the ICBS), 7,936 ineligible dwellings,[13] and 3,110 uncontacted dwellings (i.e., there was never an attempt to contact them). Since our study focuses on matched administrative records we exclude dwellings in the subset of locations that do not provide personal ID information for dwellings, which we consider

[13]Including 4,835 vacant dwellings, 1,615 dwellings that are used only occasionally, 1,065 non-residential dwellings, 22 dwellings whose tenants do not belong to the survey population, and 399 "double dwellings," i.e., single dwellings that were accidentally counted as two.

Figure 1: LFS—Sample Preparation



"nonmatchable locations."[14] The remainder forms our LFS sample of 89,152 dwellings.

The "LFS Survey Sample" (box I in Figure 1) comprises 70,581 dwellings (containing 168,159 household members) that have survey data. Another 421 dwellings are marked as respondents but lack survey data for unknown reasons. The remaining 18,150 dwellings are nonrespondents—a nonresponse rate of 20.4 percent within our LFS sample.

The "LFS Administrative Matched Sample" (box II) includes 83,638 dwellings (66,721 respondents and 16,917 nonrespondents) matched with administrative records—a high, 93.8 percent, successful match rate within our LFS sample (the match rate for nonrespondents is similarly high at 93.2 percent; see Appendix Table A.1). The remaining 5,514 dwellings could not be matched due to ID issues.[15] Appendix A.1 uses survey data to compare matched and unmatched respondents, thus assessing the implications

[14]These include localities that lack information about property taxpayers (e.g., very small and collective settlements) and certain types of buildings where not every unit has a registered property-tax payer (e.g., student dorms, assisted living, and immigration absorption centers).

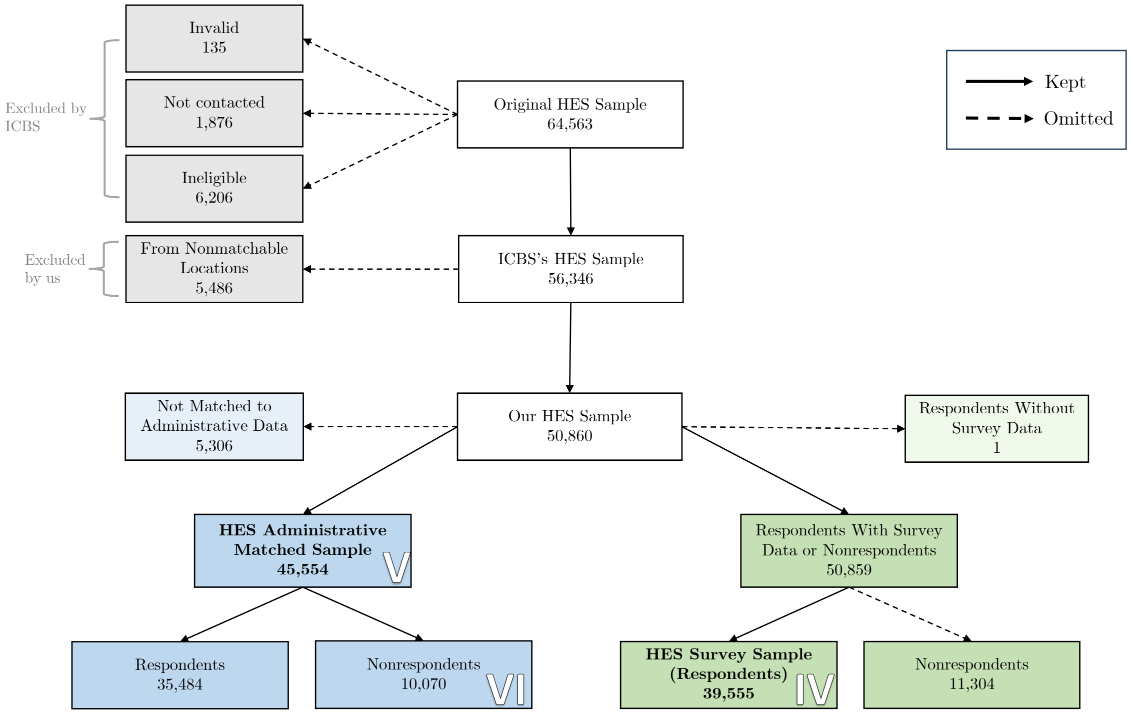
[15]Based on our conversations with the ICBS, the primary reason for unmatched ID is registration problems with local authorities, resulting in bad or missing ID data.

of excluding the unmatched dwellings. It indicates a minimal impact on our findings, mainly due to the high overall match rate.

1.3.2 HES Samples

Figure 2 depicts the HES sample creation. The initial HES sample includes 64,563 dwellings. After excluding 135 invalid, 6,206 ineligible, 1,876 uncontacted dwellings, and 5,486 dwellings from nonmatchable locations, the sample consists of 50,860 dwellings.[16]

Figure 2: HES—Sample Preparation



The "HES Survey Sample" (box IV in Figure 1) contains 39,555 dwellings with survey data. One dwelling is marked as a respondent but lacks survey data. The remaining 11,304 dwellings are nonrespondents—a nonresponse rate of 22.2 percent.

The "HES Administrative Matched Sample" (box V) comprises 45,554 dwellings (35,484 respondents and 10,070 nonrespondents) matched with administrative records,

[16]The ineligible dwellings in the HES consist of 2,830 vacant dwellings, 1,560 dwellings that are used only occasionally, 848 non-residential dwellings, 491 dwellings whose tenants do not belong to the survey population, 175 double dwellings, and 302 dwellings that are classified as ineligible by the ICBS for other reasons.

reflecting an 89.6 percent match rate (the match rate for nonrespondents is similar at 89.1 percent; see Appendix Table A.1). The analysis in Appendix A.1 indicates that the exclusion of unmatched dwellings also has a negligible effect on the HES results.

2 Respondents by Paradata

2.1 Survey Trends by Number of Contact Attempts

Following the approach of Heffetz and Reeves (2019), this section examines differences among survey respondents, categorized based on the number of attempts required to contact them. The results are presented in Table 1, whose structure will be repeated throughout the subsequent sections.

Table 1 is divided into six panels: A, B, and C provide information for the 168,159 individuals from the "LFS Survey Sample" (box I in Figure 1), while D, E, and F provide information for the 39,555 dwellings from the "HES Survey Sample" (box IV in Figure 2). The table's first three columns classify these samples into three difficulty-to-reach groups based on the number of attempts made to reach them: 1, 2, and 3 or more.

The LFS and HES differ in their survey methodologies and target populations (see Section 1.1). This allows us to assess whether the observed patterns are consistent across the surveys, despite their differences in methodologies and outcomes. Note that due to the differences between the two surveys, we compare *trends* between them, but avoid direct comparisons of *levels*.

Panels A and D reveal significant demographic trends across difficulty-to-reach groups, consistent with those in the literature. The average age decreases monotonically from the easiest-to-reach group to the hardest-to-reach group (from 44.1 to 43.3 for the LFS sample and from 49.1 to 47.3 for the HES sample). This trend is primarily driven by differences in the proportion of individuals aged 65 or more. Additionally, the average years of education increases from the easiest-to-reach to the hardest-to-reach respondents (from 12.8 to 13.6 for the LFS and from 13.1 to 14.1 for the HES), as well as the Jewish share (from 69.6 to 87.5 percent for the LFS and from 74.2 to 88.4 percent for

Table 1: Survey Results—Patterns by Difficulty to Reach

Attempts		1	2	3+	All
			LFS Survey Data		
			A. Demographics		
Age (mean)		44.1	43.7	43.3	43.8
		(0.1)	(0.1)	(0.1)	(0.0)
Years of Education (mean)		12.8	13.4	13.6	13.2
		(0.0)	(0.0)	(0.0)	(0.0)
Female (%)		52.1	52.6	53.0	52.5
		(0.2)	(0.3)	(0.2)	(0.1)
Jewish (%)		69.6	80.1	87.5	76.9
		(0.2)	(0.2)	(0.2)	(0.1)
Contacted by Phone (%)		0.2	6.7	44.5	13.6
		(0.0)	(0.1)	(0.2)	(0.1)
Number of Attempts (median)		1	2	4	2
			B. Outcomes (Unadjusted Means)		
L.F.P.:	Employed	54.6	59.9	63.9	58.3
		(0.2)	(0.2)	(0.2)	(0.1)
	Unemployed	3.6	3.6	3.6	3.6
		(0.1)	(0.1)	(0.1)	(0.0)
	Not in the Labor Force	41.8	36.4	32.5	38.0
		(0.2)	(0.2)	(0.2)	(0.1)
Labor Force Participation		58.2	63.6	67.5	62.0
		(0.2)	(0.2)	(0.2)	(0.1)
Unemployment Rate		6.2	5.7	5.4	5.8
		(0.1)	(0.1)	(0.1)	(0.1)
			C. Outcomes (Adjusted Means)		
Labor Force Participation		61.2	61.7	63.6	
		(0.4)	(0.2)	(0.2)	
Unemployment Rate		5.8	5.8	5.7	
		(0.4)	(0.2)	(0.2)	
			HES Survey Data		
			D. Demographics		
Age (mean)		49.1	48.9	47.3	48.4
		(0.1)	(0.2)	(0.1)	(0.1)
Years of Education (mean)		13.1	13.9	14.1	13.7
		(0.0)	(0.0)	(0.0)	(0.0)
Jewish (%)		74.2	84.6	88.4	82.0
		(0.4)	(0.4)	(0.3)	(0.2)
Number of Attempts (median)		1	2	4	2
			E. Outcomes (Unadjusted Means)		
Inc.:	Log NIS	9.33	9.46	9.46	9.41
		(0.01)	(0.01)	(0.01)	(0.00)
	Converted to NIS	11,293	12,823	12,830	12,216
		(87)	(120)	(99)	(58)
Exp.:	Log NIS	9.09	9.17	9.14	9.13
		(0.01)	(0.01)	(0.01)	(0.00)
	Converted to NIS	8,828	9,564	9,330	9,194
		(52)	(66)	(56)	(33)
			F. Outcomes (Adjusted Means)		
Inc.:	Log NIS	9.38	9.42	9.44	
		(0.01)	(0.01)	(0.01)	
	Converted to NIS	11,815	12,314	12,560	
		(77)	(95)	(87)	
Exp.:	Log NIS	9.12	9.14	9.13	
		(0.00)	(0.01)	(0.00)	
	Converted to NIS	9,094	9,278	9,228	
		(39)	(47)	(43)	
Obs.:	LFS (individuals)	83,163	39,681	45,315	168,159
	LFS (dwellings)	33,467	16,936	20,178	70,581
	HES (dwellings)	15,166	10,481	13,908	39,555

Notes: Samples: The LFS Survey Sample (box I in Figure 1) and HES Survey Sample (box IV in Figure 2).

the HES). Another notable trend in Panel A is the dramatic increase in the percent of participants contacted by phone across the difficulty-to-reach groups (from 0.2 to 44.5 percent). This trend is largely due to the fact that interviewers typically resort to the phone only after making several unsuccessful in-person attempts.

Panels B and E demonstrate significant trends in our outcomes of interest. Panel B shows a clear trend in labor force composition across the difficulty groups. The share of employed individuals increases monotonically from 54.6 [SE 0.2] to 63.9 [0.2] percent, the labor force participation rate rises from 58.2 [0.2] to 67.5 [0.2] percent, and the unemployment rate decreases from 6.2 [0.1] to 5.4 [0.1] percent. Panel E reveals that the average log monthly income of the easiest-to-reach respondents, after conversion back to levels, is 11,293 [87] New Israeli Shekels (NIS), which is lower than that of the hardest-to-reach (12,830 [99] NIS).[17,18] The mean log monthly expenditures also increase from the easiest-to-reach (8,828 [52] NIS after conversion) to the hardest-to-reach (9,330 [56] NIS) groups, while the 2-attempt group's expenditure (9,564 [66] NIS) is the highest.

This raises the question: Can the observed differences in outcomes across difficulty-to-reach groups be solely attributed to the observed demographic differences among them? Panels C and F address this question by presenting the outcomes adjusted for these demographic differences. Intuitively, the adjusted means for each outcome is calculated as follows.[19] First, we regress the outcomes on a set of categorical demographic variables. This set includes categorical indicators for all the demographics mentioned in Panels A (for LFS outcomes) and D (for HES outcomes), along with 2-category indicators for orthodox and for having children in the household, as well as 5-category indicators for marital status and household size.[20] Additionally, we incorporate interaction terms between each demographic variable and the difficulty-to-reach indicator to capture po-

[17]During the period of 2012–2016, the NIS was worth around \$0.25–\$0.30.

[18]All analyses concerning income and expenditures in this study use the log of (1 + NIS), with reported figures converted back to NIS using an exponential of the mean log. The standard errors were transformed using the delta method.

[19]Here, we use the STATA 16.1 command "margins," which computes adjusted means and their accompanying standard errors. The explanation of the adjusted means is informed by Williams (2011).

[20]For LFS outcomes, we also include 2-category indicators for female orthodox, female Jew, and female Arab. For the calculation of the adjusted mean log expenditures, we also include a categorical income indicator.

tential heterogeneity of the effects across groups. Using the regression coefficients, we calculate the predicted outcome for each observation in our sample three times, based on the observation's real demographics and an artificial difficulty-to-reach indicator, which is set each time to a different level (1, 2, and 3+). This process generates three hypothetical samples, one for each difficulty group, each containing all the observations in our dataset. The three adjusted means represent the average predicted outcome for these three hypothetical samples. Thus, the adjusted means for each difficulty-to-reach group are calculated based on the same sample and, therefore, the same demographics.

Panels C and F reveal that even after accounting for observable demographic differences, some significant outcome differences persist, albeit reduced substantially compared to the unadjusted means. Specifically, the adjusted labor force participation rates of the 1-attempt and 2-attempts groups (61.2 [SE 0.4] and 61.7 [0.2] percent, respectively) are significantly lower than that of the hardest-to-reach group (63.6 [0.2] percent) in Panel C. In contrast, there are no statistically significant differences in the adjusted unemployment rate of the three contact groups. In Panel F, there is a monotonic increasing trend in the adjusted mean log income (from 11,815 [77] to 12,560 [87] NIS after conversion). Additionally, there is a small (but statistically significant) increase in the adjusted mean log expenditures from the easiest-to-reach group (9,094 [39] NIS) to the hardest-to-reach group (9,228 [43] NIS), while the second contact group (9,278 [47] NIS) exhibits a similar level to the hardest-to-reach group.

2.2 Administrative Trends by Number of Contacts Attempts

In Table 2, we replicate the analysis from Table 1 using administrative records—utilizing the LFS Administrative Matched Sample (Panels A–C) and the HES Administrative Matched Sample (Panels D–F) (see Section 1.3). Our goal is to verify that the patterns observed in Section 2.1 are consistent within the administrative data.[21] As detailed in Section 1.2.2, we employ three administrative proxies to represent the survey outcomes:

[21] Recall that for each dwelling, the matched administrative information pertains to the property-tax payer, rather than each household member (as in the LFS data), the household head (as in the HES demographic data), or the total household (as the HES income and expenditure data).

the “any reported income” and “full annual employment” serve as proxies for the LFS’s labor force participation measure, while individuals’ administratively recorded income serves as a proxy to the HES’s income measure.

The results in the table confirm that the patterns observed in the survey data are consistent with the administrative records. Panels A and D indicate that the easiest-to-reach respondents are older, less educated, and include a smaller proportion of Jewish individuals compared to the hardest-to-reach respondents. Additionally, all outcomes increase from the easiest-to-reach to the hardest-to-reach respondents. This trend is evident in both the unadjusted means (Panels B and E) and the adjusted means (Panels C and F), although, as before, the adjusted trends are considerably smaller when accounting for demographic differences.[22]

3 Nonrespondents by Paradata

3.1 Administrative Trends by Nonresponse Reason

In this section, we use administrative data to compare the demographics and outcomes of respondents by difficulty of reaching with nonrespondents by subgroups. As discussed in Section 1.1, the paradata of both the LFS and HES categorizes nonrespondents based on their reasons for nonresponse. Table 3 presents the distribution of nonrespondent dwellings across nonresponse categories for both surveys.

The “Absent” category refers to dwellings whose members were temporarily not at home during interview attempts.[23] “Refused” denotes households that explicitly declined participation. “Temporary Difficulty” contains households facing short-term impediments, like illness or time constraints. “Language Difficulty” describes households facing language barriers to survey completion, often Jewish immigrants from the former Soviet Union or Ethiopia. “Permanent Difficulty” applies to dwellings with mem-

[22]The adjusted means in this section follow the same calculation method as described in Section 2.1. For the computation, we used the demographic variables from Panels A and D, as well as binary indicators for female Jews and female Arabs.

[23]In the LFS, the Absent category also covers nonrespondents who did not respond to phone contacts.

Table 2: Administrative Results—Respondents

Attempts		1	2	3+	All
			LFS Administrative Matched Sample		
			A. Demographics		
Age (mean)		57.1	55.7	54.9	56.1
		(0.1)	(0.1)	(0.1)	(0.1)
Years of Education (mean)		12.4	12.9	13.0	12.7
		(0.0)	(0.0)	(0.0)	(0.0)
Female (%)		28.7	30.5	32.7	30.3
		(0.3)	(0.4)	(0.3)	(0.2)
Jewish (%)		75.6	84.1	89.3	81.6
		(0.2)	(0.3)	(0.2)	(0.2)
Number of Attempts (median)		1	2	4	2
			B. Outcomes (Unadjusted Means)		
L.F.P:	Any Reported Income	58.0	61.4	64.4	60.7
		(0.3)	(0.4)	(0.3)	(0.2)
	Full Annual Employment:	39.7	42.5	44.9	41.9
		(0.3)	(0.4)	(0.4)	(0.2)
			C. Outcomes (Adjusted Means)		
L.F.P:	Any Reported Income:	60.1	60.5	62.5	60.8
		(0.2)	(0.3)	(0.3)	(0.2)
	Full Annual Employment:	41.6	41.7	43.1	42.0
		(0.3)	(0.3)	(0.3)	(0.2)
			HES Administrative Matched Sample		
			D. Demographics		
Age (mean)		56.4	55.7	54.7	55.6
		(0.2)	(0.2)	(0.2)	(0.1)
Years of Education (mean)		12.5	13.0	13.0	12.8
		(0.0)	(0.0)	(0.0)	(0.0)
Female (%)		28.0	29.1	31.0	29.4
		(0.4)	(0.5)	(0.4)	(0.2)
Jewish (%)		77.2	85.7	89.2	83.8
		(0.4)	(0.4)	(0.3)	(0.2)
Number of Attempts (median)		1	2	4	2
			E. Outcomes (Unadjusted Means)		
Inc.:	Log NIS:	8.70	8.81	8.81	8.77
		(0.02)	(0.02)	(0.02)	(0.01)
	Converted to NIS:	6,026	6,687	6,680	6,439
		(120)	(151)	(131)	(76)
			F. Outcomes (Adjusted Means)		
Inc.:	Log NIS:	8.72	8.77	8.79	8.76
		(0.02)	(0.02)	(0.02)	(0.01)
	Converted to NIS:	6,139	6,434	6,590	6,351
		(116)	(139)	(124)	(71)
Obs.:	LFS (dwellings)	31,411	16,039	19,271	66,721
	HES (dwellings)	13,360	9,483	12,641	35,484

Notes: Sample: The respondents from the LFS Administrative Matched Sample (box II in Figure 1) and the HES Administrative Matched Sample (box V in Figure 2). The adjusted and unadjusted means for income are calculated for the 28,309 dwellings that were matched to the Individual Income Registry. Nonrespondents are included in the calculation of the adjusted means.

Table 3: Reasons for Nonresponse

NR Reason	LFS Frequency	LFS Percent	HES Frequency	HES Percent
Absent	7,540	44.9	1,213	12.0
Refused to Answer	3,836	22.8	3,878	38.5
Temporary Difficulty	2,998	17.8	570	5.7
Language Difficulty	892	5.3	215	2.1
Permanent Difficulty	203	1.2	671	6.7
Unlocated	399	2.4	339	3.4
Withdrew	0	0.0	3,122	31.0
Other	933	5.6	62	0.6
Total	16,801	100.0	10,070	100.0

Notes: Samples: The nonrespondents from the LFS Administrative Matched Sample (box III in Figure 1) and the HES Administrative Matched Sample (box VI in Figure 2). In the LFS sample, 116 nonrespondents have missing values for their nonresponse reason.

bers unable to respond due to long-standing illness or disability. "Unlocated" refers to dwellings that the interviewers were unable to locate. Finally, the "Withdrew" category, which only appears in the HES, captures households that responded to the first stage's questionnaire but withdrew before or during the diary stage. As Table 3 shows, these categories contain 94.4 percent of the LFS nonrespondents and 99.4 percent of the HES nonrespondents. The remaining nonrespondents are classified as "Other."

A comparison of the HES and LFS reveals that the LFS contains significantly higher shares of Absent and Temporary nonrespondents and a notably lower share of Refused nonrespondents. According to ICBS personnel, these differences may stem from varying survey methodologies: the HES requires a commitment upfront for a relatively demanding process (including the diary-keeping process), which could lead to higher refusal rates, while the LFS's limited time frame for follow-up attempts may restrict opportunities to connect with temporarily unavailable nonrespondents. These differences, along with the large but HES-specific category of Withdrew, underscore how survey methodology can significantly influence the composition of nonrespondents (which, as demonstrated in this section, is a key factor in determining the nature and extent of nonresponse bias).

The hypothesis that nonrespondents resemble hard-to-reach respondents is rooted in the idea that nonrespondents are essentially respondents who just require additional attempts to be contacted. This hypothesis seems plausible for nonrespondents from the Absent and Temporary categories, who may indeed have only needed more contact attempts to be reached. However, it seems less applicable to those in the Permanent, Language, Refused, Withdrew, and Unlocated categories. The inability of these nonrespondents to participate due to language barriers, permanent issues, unwillingness to engage, a decision to withdraw mid-survey, or being unlocatable suggests that they differ from typical respondents in ways beyond just needing more contact attempts.

The results in Table 4, presenting the demographics and outcomes by nonresponse categories, largely support these conjectures. Specifically, the Absent nonrespondents resemble hard-to-reach respondents in most demographics and outcomes (see Table 2 for respondents). Specifically, they align with the hardest-to-reach group in terms of average age and education. Its percent Jewish resembles the hardest-to-reach respondents in the LFS and the 2-attempts respondents in the HES. Before adjusting for demographic differences, this category's outcomes—the "any reported income" (64.8 [SE 0.5] percent), "full annual employment" (44.6 [0.6] percent), and log income (6,501 [397] NIS after conversion) are similar to those of the hardest-to-reach respondents (64.4 [0.3], 44.9 [0.3], and 6,680 [131], respectively). These outcome similarities remain also after demographic adjustments (63.4 [0.5] vs. 62.5 [0.3] percent, 43.2 [0.5] vs. 43.1 [0.3] percent, and 6,697 [406] vs. 6,590 [124] NIS, respectively).

The Temporary Difficulty category also exhibits a "hard-to-reach" profile. These nonrespondents are younger and have similar education attainments compared to the hardest-to-reach respondents. Its percent Jewish resembles the hardest-to-reach respondents in the LFS, but is akin to that of the easiest to reach in the HES. Before adjustments, its labor force participation ("any reported income" of 65.2 [0.9] and "full annual employment" of 44.3 [0.9]) and income (6,985 [526] NIS) are similar to those of the hardest-to-reach respondents. After adjustments, the "any reported income" (61.5 [0.8] percent) falls between the 2-attempts (60.5 [0.3]) and 3+-attempts (62.5 [0.3]) respondents, the "full annual employment" (41.2 [0.8] percent) is slightly below the easiest-to-reach respondents (41.6 [0.3]), and the income (6,724 [591] NIS) slightly exceeds that of the hardest-to-reach respondents (6,590 [124]).

Table 4: Nonrespondents by Reason

		Absent	Refused	Temporary	Language	Permanent	Unlocated	Withdraw	Other
					LFS Administrative Matched Sample				
					A. Demographics				
Age (mean)		54.1	56.8	53.0	61.0	68.0	56.3		53.0
		(0.2)	(0.3)	(0.3)	(0.6)	(1.5)	(1.1)		(0.6)
Years of Education (mean)		13.0	12.3	12.8	11.8	10.1	11.1		12.2
		(0.0)	(0.1)	(0.1)	(0.2)	(0.4)	(0.2)		(0.1)
Female (%)		33.9	32.7	29.7	37.4	45.8	38.6		28.9
		(0.5)	(0.8)	(0.8)	(1.6)	(3.5)	(2.4)		(1.5)
Jewish (%)		88.9	87.9	88.0	80.5	81.3	52.9		66.9
		(0.4)	(0.5)	(0.6)	(1.3)	(2.7)	(2.5)		(1.5)
Attempts (median)		5	4	5	3	2	2		3
					B. Outcomes (Unadjusted Means)				
L.F.P:	Reported Income	64.8	59.0	65.2	52.1	24.6	48.4		63.3
		(0.5)	(0.8)	(0.9)	(1.7)	(3.0)	(2.5)		(1.6)
	Full Annual Emp.	44.6	39.0	44.3	36.4	16.7	30.3		43.5
		(0.6)	(0.8)	(0.9)	(1.6)	(2.6)	(2.3)		(1.6)
					C. Outcomes (Adjusted Means)				
L.F.P:	Reported Income	63.4	61.1	61.5	60.5	48.0	54.7		62.2
		(0.5)	(0.7)	(0.8)	(1.6)	(4.0)	(2.6)		(1.4)
	Full Annual Emp.	43.2	40.9	41.2	42.4	35.3	37.5		43.3
		(0.5)	(0.8)	(0.8)	(1.8)	(4.4)	(2.8)		(1.6)
					HES Administrative Matched Sample				
					D. Demographics				
Age (mean)		54.9	55.8	53.0	63.6	65.2	55.9	56.3	56.1
		(0.5)	(0.3)	(0.7)	(1.3)	(0.8)	(1.0)	(0.3)	(2.4)
Years of Education (mean)		13.1	12.6	13.0	11.9	11.9	12.6	12.2	12.0
		(0.1)	(0.1)	(0.1)	(0.4)	(0.2)	(0.2)	(0.1)	(0.4)
Female (%)		35.6	32.2	28.2	39.5	42.6	34.8	30.7	22.6
		(1.4)	(0.8)	(1.9)	(3.3)	(1.9)	(2.6)	(0.8)	(5.4)
Jewish (%)		84.7	85.7	77.4	91.6	92.0	83.5	82.2	75.8
		(1.0)	(0.6)	(1.8)	(1.9)	(1.1)	(2.0)	(0.7)	(5.5)
Attempts (median)		5	4	4	2	2	2	2	2
					E. Outcomes (Unadjusted Means)				
Inc.:	Log NIS	8.78	8.73	8.85	8.50	8.34	8.55	8.59	8.82
		(0.06)	(0.04)	(0.08)	(0.18)	(0.15)	(0.13)	(0.04)	(0.26)
	Converted to NIS	6,501	6,190	6,985	4,934	4,192	5,180	5,353	6,774
		(397)	(220)	(526)	(903)	(622)	(663)	(232)	(1,742)
					F. Outcomes (Adjusted Means)				
Inc.:	Log NIS	8.81	8.76	8.81	8.79	8.52	8.62	8.65	8.51
		(0.06)	(0.03)	(0.09)	(0.27)	(0.12)	(0.12)	(0.04)	(0.35)
	Converted to NIS	6,697	6,400	6,724	6,564	5,001	5,547	5,730	4,952
		(406)	(218)	(591)	(1,771)	(606)	(669)	(222)	(1,719)
Obs.:	LFS sample	7,540	3,836	2,998	892	203	399		933
	HES sample	1,213	3,878	570	215	671	339	3,122	62

Notes: Sample: the LFS and HES administrative matched samples. In the LFS sample, 116 dwellings are missing nonresponse reason data and are therefore excluded from the analyses in this section. The adjusted and unadjusted means income are calculated for the 28,309 dwellings that have reported income. Respondents are included in the calculation of the adjusted means.

Conversely, the Refused category resembles the easiest-to-reach respondents in most metrics. Its age is situated between the easiest-to-reach and 2-attempt respondents, and its education level is similar to that of the easiest-to-reach group. Its percent Jewish falls between the 2-attempts and 3+-attempts groups in the LFS, and it is similar to the 2-attempts group in the HES. Additionally, before adjustments, its labor force participation ("any reported income" of 59.0 [0.8] and "full annual employment" of 39.0 [0.8]) and income (6,190 [220] NIS) are most similar to those of the easiest-to-reach respondents (58.0, 39.7, and 6,026, respectively). After adjustments, its "any reported income" (61.1 [0.7] percent) falls between the 2-attempts (60.5 [0.3]) and 3+-attempts (62.5 [0.3]) groups, its "full annual employment" (40.9 [0.8] percent) is lower than that of the easiest-to-reach group (41.6 [0.3]), and income (6,400 [218] NIS) is similar to that of the 2-attempts group (6,434 [139]).

The Language Difficulty nonrespondents are considerably older and less educated than even the easiest-to-reach respondents. Moreover, all their unadjusted outcomes—the "any reported income" (52.1 [1.7] percent), "full annual employment" (36.4 [1.6] percent), and income (4,934 [903] NIS after conversion)—are considerably lower than those of the easiest-to-reach respondents. Interestingly, adjusting for observable demographic differences elevates them to fall within the respondents' range.

The Permanent Difficulty category is also significantly older and less educated than the easiest-to-reach respondents. Furthermore, its unadjusted "any reported income," "full annual employment," and income are considerably lower than even those of the Language category. Adjusting for observable demographic differences leads to a substantial increase in these outcomes. However, even after this increase, all outcomes—the "any reported income" (48.0 [4.0] percent), "full annual employment" (35.3 [4.4] percent), and income (5,001 [606])—remain considerably lower than the easiest-to-reach respondents (60.1 [0.2], 41.6 [0.3], 6,139 [116]).

The Unlocated category also generally exhibits an "easy-to-reach" demographic profile. Additionally, its unadjusted "any reported income" (48.4 percent [2.5]) and "full annual employment" (30.3 percent [2.3]) are lower, and its income (5,180 [663] NIS) is similar to the Language Difficulty category. Adjusting for demographics increases this category's outcomes as well, but they remain considerably lower than even those of the easiest-to-reach respondents. However, standard errors are large.

The Withdrew category, which exists only in the HES, exhibits a similar age and lower education level compared to the easiest-to-reach respondents. Its unadjusted and adjusted income (5,353 [232] and 5,730 [222] NIS, respectively) is lower than that of the easiest-to-reach respondents (6,026 [120] and 6,139 [116]).

Finally, in the LFS, the Other category's age is lower than even the hardest-to-reach respondents, but its education level and percentage of Jewish individuals are lower than those of the easiest-to-reach respondents. Both its "any reported income" (63.3 percent [1.6] before adjustment and 62.2 percent [1.4] after) and "full annual employment" (43.5 percent [1.6] before adjustment and 43.3 percent [1.6] after) are similar to those of the hardest-to-reach respondents. We do not separately analyze the Other category in the HES as it only contains 62 dwellings.

One might wonder about further heterogeneity within each nonresponse category. A high level of heterogeneity within a particular category could indicate that this subgroup comprises multiple distinct subgroups, suggesting that further categorization could be useful. To explore this question, Appendix Table A.3 presents the standard deviation of the non-binary variables—age, years of education, and income—within each nonresponse category and for the entire nonrespondent sample. Interestingly, the smaller categories—Language Difficulty, Permanent Difficulty, and Unlocated—exhibit greater heterogeneity, while the larger categories—Absent, Refused, and Temporary—are more homogeneous, and are slightly more homogenous than the entire nonrespondent sample.

In summary, classifying nonrespondents by their nonresponse reasons reveals that they are a heterogeneous group with diverse demographics and outcomes. While some nonrespondents indeed fit on the "hard-to-reach" side of the difficulty-of-reaching scale, others are actually more similar to the easy-to-reach respondents. Overall, these results reveal a strong relationship between nonresponse reason and the nonrespondent's profile. As discussed in Section 4, this relationship can be helpful in assessing potential

Table 5: Number of Contact Attempts by Nonresponse Categories

Attempts	1	2	3	4	5	6	7	8+	Total
					A. LFS				
Absent	23.4	33.3	41.6	47.2	50.9	53.8	52.9	46.7	44.9
Refused to Answer	17.9	23.5	24.5	24.6	23.2	20.1	20.7	23.7	22.8
Temporary Difficulty	18.0	16.2	15.9	14.9	15.3	17.6	21.5	25.0	17.8
Language Difficulty	2.6	10.8	8.0	6.6	5.1	4.0	1.6	0.9	5.3
Permanent Difficulty	4.6	2.8	1.0	1.1	0.8	0.5	0.3	0.4	1.2
Unlocated + Other	33.6	13.5	9.0	5.7	4.6	4.0	3.0	3.3	7.9
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
					B. HES				
Absent	3.6	6.0	8.6	13.3	20.4	16.5	23.8	26.6	12.0
Refused to Answer	17.1	32.1	42.2	47.1	50.4	43.9	51.4	52.1	38.5
Temporary Difficulty	3.0	5.5	6.1	6.3	7.6	4.5	5.9	8.0	5.7
Language Difficulty	3.1	2.6	2.5	2.2	1.4	0.2	1.3	1.2	2.1
Permanent Difficulty	11.6	8.8	6.0	5.6	3.8	5.3	1.5	1.4	6.7
Withdrew	54.9	39.5	30.2	21.8	13.8	28.8	15.1	10.5	31.0
Unlocated + Other	6.7	5.6	4.4	3.6	2.4	0.8	1.0	0.1	4.0
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Samples: The nonrespondents from the LFS Administrative Matched Sample (box III in Figure 1) and the HES Administrative Matched Sample (box VI in Figure 2). For the Withdrew category, the number of contact attempts refers to the count of attempts made before the initial successful contact.

nonresponse bias in survey outcomes.

3.2 Administrative Trends by Number of Contact Attempts

We now classify nonrespondents by number of (ultimately unsuccessful) contact attempts. This classification offers an alternative approach for surveys that track contact attempts but do not record the reasons for nonresponse.

Interviewers may make more such attempts to nonrespondents whom they believe are more likely to be converted into respondents with additional contact attempts. This conjecture is supported by Table 5, which presents the distribution of nonresponse reasons across contact groups. It shows that as the number of contact attempts increases, the proportion of Absent and Temporary categories rises, whereas the share of Language and Permanent categories decreases. The conjecture implies that nonrespondents with more contact attempts should exhibit a “harder-to-reach” profile.

Table 6 supports this hypothesis. Panels A and D reveal that those approached fewer times tend to be older, less educated, have a lower percent Jewish, participate less in the labor force, and have lower incomes compared to those with more attempted contacts. Panels B and C display significant upward trends in the "any reported income" (from 52.7 [SE 1.6] to 67.5 [0.9] percent for unadjusted means and 56.9 [1.4] to 64.5 [0.9] percent for adjusted means) and "full annual employment" (from 35.4 [1.5] to 47.7 [1.0] for unadjusted means and from 39.6 [1.6] to 44.8 [1.0] percent for adjusted means) proxies as the number of contact attempts increases. Panels E and F demonstrate an increasing trend in the log income from the 1-attempt group (5,374 [300] and 5,699 [308] NIS for unadjusted and adjusted means) to the 5-attempts group (6,349 [354], 6,623 [366] NIS), but a noticeable decline in the 7-attempts (5,420 [651], 5,562 [674] NIS) and 8+-attempts (5,426 [473], 5,885 [498] NIS) groups. The reasons behind this decrease (which is not statistically significant) are not clear to us.

3.3 Nonresponse Bias in Administrative Outcomes

Table 7 presents the nonresponse bias in the administrative outcomes. Column 1 shows the (unadjusted) average outcomes for all nonrespondents, while Column 2 displays the average for respondents (as detailed in Table 2). Column 3 presents the average for all dwellings, representing an average of nonrespondent and respondent outcomes weighted by the response rate. Finally, Column 4 calculates the nonresponse bias, i.e., the percentage by which the average outcomes of respondents differ from that of the entire sample. The results indicate a limited nonresponse bias in all outcomes: -0.4 percent in the "any reported income" proxy, -0.1 percent in the "full annual employment" proxy, and +2.1 percent in log income (after it is converted back to NIS).[24]

[24]To put this into perspective, between 1990 and 2022, the labor force participation rate for individuals aged 15 and older in Israel ranged from a low of 59.1 percent in 1990 to a high of 64.2 percent in 2014, with a standard deviation of 1.3 percentage points during this period (OECD (2024)).

Table 6: Administrative Results—Nonrespondents by Contact Attempts

Attempts		1	2	3	4	5	6	7	8+
					LFS Administrative Matched Sample				
					A. Demographics				
Age (mean)		55.2	56.6	56.3	55.4	55.0	54.7	53.5	53.2
		(0.6)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.3)
Female (%)		30.4	33.1	32.6	34.1	33.3	32.0	34.3	32.9
		(1.4)	(1.0)	(0.9)	(0.9)	(1.0)	(1.0)	(1.2)	(0.9)
Years of Education (mean)		11.5	12.3	12.5	12.8	12.6	12.8	12.9	12.9
		(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
Jewish (%)		53.5	78.8	84.9	88.3	90.4	92.1	92.2	90.6
		(1.6)	(0.9)	(0.7)	(0.6)	(0.6)	(0.6)	(0.7)	(0.6)
Number of Attempts (median)		1	2	3	4	5	6	7	9
					B. Outcomes (Unadjusted Means)				
L.F.P:	Any Reported Income	52.7	56.1	59.6	61.7	62.9	63.9	67.3	67.5
		(1.6)	(1.1)	(0.9)	(0.9)	(1.0)	(1.1)	(1.2)	(0.9)
	Full Annual Employment:	35.4	37.5	38.4	42.6	43.8	42.4	46.3	47.7
		(1.5)	(1.1)	(0.9)	(1.0)	(1.0)	(1.1)	(1.3)	(1.0)
					C. Outcomes (Adjusted Means)				
L.F.P:	Any Reported Income:	56.9	59.0	61.2	61.9	62.8	62.5	64.8	64.5
		(1.4)	(0.9)	(0.8)	(0.8)	(0.9)	(1.0)	(1.2)	(0.9)
	Full Annual Employment:	39.6	39.7	39.7	42.5	43.5	41.1	43.5	44.8
		(1.6)	(1.0)	(0.9)	(0.9)	(1.0)	(1.1)	(1.3)	(1.0)
					HES Administrative Matched Sample				
					D. Demographics				
Age (mean)		58.5	57.1	56.5	56.4	55.7	55.4	53.8	54.0
		(0.4)	(0.4)	(0.5)	(0.5)	(0.5)	(0.7)	(0.8)	(0.6)
Years of Education (mean)		11.9	12.5	12.5	12.5	12.9	12.7	12.8	12.7
		(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.2)	(0.1)
Female (%)		30.1	30.4	33.0	33.5	36.5	36.3	31.7	34.4
		(1.0)	(1.1)	(1.2)	(1.3)	(1.2)	(1.9)	(2.4)	(1.7)
Jewish (%)		74.4	80.8	86.0	88.5	89.8	88.5	91.3	91.9
		(1.0)	(0.9)	(0.9)	(0.9)	(0.8)	(1.3)	(1.4)	(1.0)
Number of Attempts (median)		1	2	3	4	5	6	7	9
					E. Outcomes (Unadjusted Means)				
Inc.:	Log NIS:	8.59	8.68	8.67	8.75	8.76	8.72	8.60	8.60
		(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.08)	(0.12)	(0.09)
	Converted to NIS:	5,374	5,860	5,833	6,312	6,349	6,101	5,420	5,426
		(300)	(312)	(343)	(408)	(354)	(501)	(651)	(473)
					F. Outcomes (Adjusted Means)				
Inc.:	Log NIS:	8.65	8.71	8.70	8.79	8.80	8.77	8.62	8.68
		(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.09)	(0.12)	(0.08)
	Converted to NIS:	5,699	6,091	6,007	6,555	6,623	6,441	5,562	5,885
		(308)	(309)	(324)	(397)	(366)	(552)	(674)	(498)
Obs.:	LFS (dwellings)	1,024	2,072	2,760	2,663	2,448	1,977	1,478	2,495
	HES (dwellings)	1,985	1,851	1,581	1,297	1,561	642	391	762

Notes: Sample: The nonrespondents from the LFS Administrative Matched Sample (box III in Figure 1) and the HES Administrative Matched Sample (box VI in Figure 2). Though the results for the respondents are not displayed, they are incorporated into the adjusted means computation, including both the calculation of the regression coefficients and the final determination of the means.

Table 7: Nonresponse Bias in Administrative Outcomes

	Nonrespondents	Respondents	All	Nonresponse Bias
		LFS Outcomes		
Any Reported Income	61.9	60.7	60.9	−0.4%
Full Annual Employment	42.1	41.9	41.9	−0.1%
		HES Outcomes		
Log Inc. (Converted to NIS)	5,854	6,439	6,309	+2.1%

Notes: This table presents the nonresponse bias in our administrative outcomes based on the matched administrative records. For a detailed description, refer to Section 3.3.

4 Assessing Nonresponse Bias in Survey Outcomes

4.1 LFS and HES Outcomes

4.1.1 Testing Our Method: Administrative Outcomes

In the remainder of the paper, we demonstrate how the insights from Section 3 can help assess nonresponse bias in survey outcomes when administrative records are not available—relying solely on the (frequently available) combination of survey data and paradata. However, before we address the survey outcomes, we illustrate our methodology for the administrative outcomes, for which the nonresponse bias is known (see Section 3.3).

The two key components for this assessment are the nonrespondents' distribution by nonresponse category (see Panel A of Table 8)[25] and the administrative outcomes of the respondents based on the difficulty of reaching them (as shown in Table 2).

Based on our comparisons of nonrespondents to respondents from Section 3.1, we extrapolate (unadjusted) outcomes for the nonrespondent categories using the outcomes of respondents by difficulty of reaching. Panel B of Table 8 presents the extrapolated administrative outcomes for each nonresponse category. Specifically, we assume that the outcomes for the Absent and Temporary Difficulty categories are similar to those of the hardest-to-reach respondents. Conversely, the outcomes for the Refused and Withdrew

[25] For the nonresponse bias assessment, we use the distribution of nonresponse reasons for all nonrespondents, not just those matched with administrative records (the distribution for the matched dwellings is presented in Table 3). However, these distributions are nearly identical.

Table 8: Extrapolating Outcomes for Nonresponse Categories

	Absent	Refused	Temporary	Language	Permanent	Unlocated	Other	Withdrew
				A. Distribution (%)				
LFS	44.9	22.8	17.8	5.3	1.2	2.4	5.6	0.0
HES	12.0	38.5	5.7	2.1	6.7	3.4		31.0
				B. Extrapolated Administrative Outcomes				
				LFS Outcomes				
Any Reported Income	64.4	58.0	64.4	54.7	51.3	54.7	64.4	
Full Annual Employ.	44.9	39.7	44.9	36.9	34.2	36.9	44.9	
				HES Outcomes				
Income: Log	8.81	8.70	8.81	8.60	8.39	8.60		8.70
NIS	6,680	6,026	6,680	5,430	4,414	5,430		6,026
				C. Extrapolated Survey Outcomes				
				LFS Outcomes				
Labor Force Particip.	67.5	58.2	67.5	52.8	47.4	52.8	67.5	
Unemployment Rate	5.4	6.2	5.4	6.7	7.1	6.7	5.4	
				HES Outcomes				
Inc.: Log	9.46	9.33	9.46	9.20	8.95	9.20		9.33
NIS	12,830	11,293	12,830	9,946	7,710	9,946		11,293
Exp.: Log	9.14	9.09	9.14	9.01	8.87	9.01		9.09
NIS	9,330	8,828	9,330	8,148	7,115	8,148		8,828

Notes: This table presents our extrapolation of the survey outcomes for the nonresponse categories based on our findings. We omitted the dwelling from the Other category in the HES due to low number of observations.

categories are assumed to be similar to those of the easiest-to-reach respondents. For the Language Difficulty and Unlocated categories, we assume their outcomes are lower than those of the easiest-to-reach respondents, extrapolated using the observed decrease from the second difficulty group to the first. For the Permanent Difficulty category, we assume outcomes are lower than those of the Language Difficulty and Unlocated categories by the same margin that these categories differ from the easiest-to-reach respondents. Finally, we assume that the LFS outcomes for the Other category are similar to those of the hardest-to-reach respondents. Due to the low number of observations, we omit the Other category dwellings from the HES analysis. While these extrapolations are merely illustrative and based on (somewhat crude) assumptions, they are arguably no more arbitrary—and they better align with the evidence provided in this paper—than simply assuming that all nonrespondents are either (conditionally) missing at random or harder to reach than the hardest-to-reach respondents.

Using these extrapolated outcomes for the nonresponse categories, we calculate av-

Table 9: Assessing Nonresponse Bias in LFS and HES Outcomes

	(1)	(2)	(3)	(4)
	Extrapolated Outcomes: All Nonrespondents	Outcomes: Respondents	Extrapolated Outcomes: Intended Population	Approximate Nonresponse Bias
		A. Administrative Outcomes		
		LFS Outcomes		
Any Reported Income	62.1	60.7	61.0	−0.5%
Full Annual Employment	43.0	41.9	42.1	−0.5%
		HES Outcomes		
Log Inc. (Converted to NIS)	5,672	6,439	6,261	+2.8%
		B. Survey Outcomes		
		LFS Outcomes		
Labor Force Participation	64.0	62.0	62.4	−0.7%
Unemployment Rate	5.7	5.8	5.8	+0.6%
		HES Outcomes		
Log Inc. (Converted to NIS)	11,180	12,216	11,978	+2.0%
Log Exp. (Converted to NIS)	8,743	9,194	9,092	+1.1%

Notes: Column 1 presents the extrapolated nonrespondent averages for the outcomes based on the extrapolations in Table 8. Column 2 presents the average outcome of the respondents, as found in Tables 2 for the administrative outcomes and 1 for the survey outcomes. Column 3 presents the extrapolated expenditures for the entire intended population by averaging columns 1 and 2, weighted by the response rate at each survey. Column 4 quantifies the percentage by which the average outcomes of the respondents exceed our population estimates.

erage outcomes for all nonrespondents, weighted by the categories' shares. Column 1 of Table 9 presents the extrapolated nonrespondent averages for our three administrative outcomes. Column 2 presents the average outcomes of the respondents, as found in Table 2. Column 3 presents the extrapolated expenditures for the entire intended population by averaging the nonrespondent extrapolated outcomes and respondent outcomes, weighted by the response rate (79.6 percent in the LFS and 77.8 percent in the HES). Finally, Column 4 assesses the nonresponse bias for each outcome: −0.5 percent for the "any reported income" proxy, −0.5 percent for the "full annual employment"' proxy, and +2.8 percent for log income (converted back to NIS). These three figures (−0.5,−0.5, and +2.8 percent) would be our extrapolation method's assessment of nonresponse bias had we applied it to our administrative records. They are on average closer to the empirically observed bias in Table 7 (−0.4, −0.1, and +2.1 percent, respectively) than the standard assumption of no bias, although for the administrative outcomes, the overall bias is admittedly limited.

4.1.2 Applying Our Method: Survey Outcomes

We move to estimating the nonresponse bias in the LFS and HES survey outcomes, for which we do not have the actual nonresponse bias values. We use the same procedure as described above, utilizing the survey outcomes of respondents by difficulty of reaching (shown in Table 1). Panel C of Table 8 presents the extrapolated survey outcomes for the nonresponse categories, based on the same assumptions presented above for the administrative outcomes. Column 4 of Panel B in Table 9 shows our estimates for the nonresponse bias in the survey outcomes. Consistent with our results for administrative outcomes, the estimates suggest a limited nonresponse bias for the survey outcomes: −0.7 percent for the labor force participation rate, +0.6 percent for the unemployment rate, +2.0 percent for log income (converted back to NIS), and +1.1 percent for (converted) log expenditures. Note that these figures are not expected to exactly match the nonresponse bias figures in the administrative outcomes (section 4.1.1) since the administrative proxies are not perfect substitutes for the survey outcomes.

4.2 U.S. Consumer Expenditure Survey (CE)

Can our empirical findings help in assessing nonresponse bias in surveys that, like most surveys, cannot be readily matched with administrative records? In this section we illustrate how our methodology can be applied in such cases, under the assumption that the similarities we found in our administrative records between paradata-based groups of nonrespondents and respondents extend beyond the two surveys we examined. At present, it is a strong assumption. We hope that future evidence from administrative records like ours will help replace it with empirical evidence.

We illustrate our adjusted methodology using the U.S. Consumer Expenditure Survey (CE) interview data from the first quarter of 2021 to the first quarter of 2023 (for files and documentation, see U.S. Bureau of Labor Statistics (2023a)). CE data underlies important macro indicators in the U.S., including household expenditures and the consumer price index (CPI, where CE data underlie the index weights). The CE

Table 10: CE—Distribution by Nonresponse Category

	Absent	Refused (Hostile)	Refused (Other)	Time Const.	Language	Other	Total
Frequency	12,549	6,885	22,001	9,303	803	4,691	56,232
Percent	22.3	12.2	39.1	16.5	1.4	8.3	100.0

Notes: Source: U.S. Bureau of Labor Statistics (2023a). See the end of Appendix A.3 for technical details, including files and variables used.

response rate has steadily declined over the years, reaching a low of 40 to 45 percent in 2023—with potentially meaningful increases in nonresponse bias.

We start with the two essential components of our approach. The first, presented in Table 10, is the distribution of nonresponse reasons among the nonrespondents as categorized by the CE. The second is the survey respondents' outcome variable—here, reported quarterly household expenditures—categorized by the difficulty of reaching them. The distribution of number of contact attempts in the CE looks rather different from that in the LFS and HES, with notably more attempts, on average, in the CE. Therefore, we propose three alternative ways to categorize respondents by difficulty of reaching: eight difficulty groups ranging from 1 to 8+ contact attempts; three difficulty groups of 1, 2, and 3+ attempts (the same categorization we used in our LFS and HES analyses); and three difficulty groups, as balanced in number of respondents as possible (1–2, 3–4, and 5+). (See Appendix A.3 for respondent distributions and average log expenditures by difficulty group for each categorization alternative.) As expected, log expenditures increase monotonically with the number of contact attempts across all categorizations. For example, they increase from $10,311 (1 attempt) to $14,024 (8+ attempts) in the eight-group categorization (Appendix Table A.4).

Next, based on our LFS and HES findings, we use respondent data to extrapolate expenditure ballpark estimates for the nonresponse categories in Table 10. We assume that the expenditures of the Refused (Other) category align with the easiest-to-reach respondents. For the Absent[26] and Time Constraint categories, we propose two alternative assumptions: either that their log expenditures align with those of the hardest-to-reach respondents or that they exceed them, continuing the increasing trend from the second

[26]The Absent category includes households categorized as "no one home" and "temporary absent."

hardest-to-reach to the hardest-to-reach group. For the Language and Refused (Hostile) categories, we also offer two alternatives: either that their log expenditures align with those of the easiest to reach, or that they are lower than them, continuing the decreasing trend from the second difficulty group to the easiest-to-reach group.[27] (These assumptions appear aligned with Appendix Table A.8, showing a significantly higher number of contact attempts for the Absent and Time Constraint categories compared to the Refused (Hostile) and Language categories, with the Refused (Other) category falling in-between.) Last, although we do not have much information about the Other category (and it is likely different across surveys), it received fewer contact attempts than even the Language and Refused (Hostile) categories (Appendix Table A.8), suggesting potential similarity with these categories.

Based on these assumptions and extrapolations, we can calculate the average log expenditures of all nonrespondents and, therefore, estimate nonresponse bias. Table 11 presents twelve alternatives for doing so. Each combines a way to categorize respondents by difficulty (as outlined in column 2) with assumptions about log expenditures for the Refused (Hostile), Language, and Other categories (column 3) and the log expenditures of the Absent and Time Constraint categories (column 4). Results suggest that average respondents' expenditures ($12,836 after converting back to levels) are 3.0 to 15.7 percent higher than those of the nonrespondents (reported in column 5), resulting in an upward nonresponse bias—relative to our estimate of average expenditures in the entire intended population (column 6)—between 1.7 to 8.4 percent (column 7).[28,29] Despite relying on simplified assumptions, this method demonstrates the potential for insights into nonresponse bias using only respondent data, supplemented with relevant paradata for both respondents and nonrespondents.

[27]The methods for extrapolating out-of-sample trends, while somewhat arbitrary, were chosen over a linear-trend prediction with respect to all difficulty categories due to the evident concavity in the relationship between contact attempts and log expenditure (see Appendix Table A.4).

[28]Nonresponse bias is calculated as the percentage by which the respondents' average expenditures (converted back to levels) differ from our population estimate.

[29]This is consistent with our HES results, where respondents' income is about 10 percent higher than that of the nonrespondents, resulting in an overall upward nonresponse bias of approximately 2 percent, as detailed in Panel E of Table 7 (the relatively low nonresponse bias is due to higher response rate in the HES than in the CE).

Table 11: Alternative Methods for Assessing Nonresponse Bias in CE Expenditure Data

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Altern-ative	Respondent Difficulty Categorization	Assumed Expenditures of Refused (Hostile), Language, and Other Relative to Easiest-to-Reach Respondents	Assumed Expenditures of Absent and Time Constraint Relative to Hardest-to-Reach Respondents	Extrapolated Expenditures: All Nonre-spondents	Extrapolated Expenditures: Intended Population	Approximate Nonresponse Bias
1	1, 2,..., 8+	Lower	Same	$11,289	$11,950	+7.4%
2	1, 2,..., 8+	Lower	Higher	$11,307	$11,960	+7.3%
3	1, 2,..., 8+	Same	Same	$11,620	$12,144	+5.7%
4	1, 2,..., 8+	Same	Higher	$11,639	$12,155	+5.6%
5	1, 2, 3+	Lower	Same	$11,096	$11,835	+8.4%
6	1, 2, 3+	Lower	Higher	$11,679	$12,178	+5.4%
7	1, 2, 3+	Same	Same	$11,422	$12,028	+6.7%
8	1, 2, 3+	Same	Higher	$12,022	$12,376	+3.7%
9	1–2, 3–4, 5+	Lower	Same	$11,807	$12,252	+4.8%
10	1–2, 3–4, 5+	Lower	Higher	$12,078	$12,408	+3.5%
11	1–2, 3–4, 5+	Same	Same	$12,179	$12,466	+3.0%
12	1–2, 3–4, 5+	Same	Higher	$12,459	$12,624	+1.7%

Notes: Each alternative for calculating nonresponse involves a unique combination of a way to categorize respondents by contact attempts (as detailed in column 2), an assumption about the log expenditures for the Refused (Hostile), Language, and Other categories (column 3), and an assumption for the Absent and Time Constraint categories (column 4) (see Appendix A.3 for average log expenditures of respondents by difficulty group for each categorization alternative). Each combination results in a slightly different extrapolation of expenditures for the nonresponse categories (see Appendix Table A.7 for the assumed log expenditures of each nonresponse category under each alternative). The nonrespondents' average log expenditures is calculated based on the extrapolated log expenditures for the nonresponse categories, weighted by their percent (column 5 reports this average after converting back to dollars). The population estimate (column 6) is derived from averaging the log expenditures of respondents (9.46) and the extrapolated log expenditures of nonrespondents, weighted by a response rate of 44.3% (and then converted back to dollars). Nonresponse bias (column 7) quantifies the percentage by which the average expenditures of the respondents ($12,836 after converting back to levels) exceed our population estimate.

We note that the CE offers an alternative approach to calculating the population average by weighting respondents using the intended population's demographics. This approach gives a demographically weighted log expenditure average of $12,758 (after converting back to levels). While this figure is 0.6 percent lower than the unweighted respondent average, it remains 1.1–7.8 percent higher than any of our population estimate alternatives. This gap could be, at least partly, because demographic differences may not completely account for the outcome differences between respondents and nonrespondents, as implied by our study's results.

5 Conclusion

The persistent decline in response rates in national surveys raises significant concern regarding the risk of nonresponse bias, including in aggregate indicators that are widely used in both economic policy and research. Estimating nonresponse bias is challenging. While administrative records may sometimes be informative, many important economic indicators cannot be readily calculated from nationally available administrative records (which is why we still rely heavily on national surveys). In this paper, we examine whether survey data and paradata can be used to evaluate this bias.

Using national Israeli registries, we match administrative records with almost the entire intended survey sample (including nonrespondents) of two prominent government surveys and construct administrative data proxies for key survey outcomes. This allows us to directly compare nonrespondents and respondents in aggregate and across the difficulty-to-reach spectrum. Our findings indicate that nonrespondents are a heterogeneous group with diverse demographics and outcomes. Those who were temporarily unavailable (due to absence or other temporary difficulties) appear as respondents who just required more attempts to be reached, i.e., their demographics and outcomes resemble those of the hardest-to-reach respondents. Conversely, nonrespondents from other subgroups—such as those refusing participation, withdrawing mid-process, or facing permanent or language barriers—appear different from typical respondents in ways beyond merely requiring more attempts to be reached. In fact, their demographics and outcomes are more similar to those of the easiest-to-reach respondents and, in some cases, appear even "easier-to-reach" than them. Furthermore, nonrespondents with fewer contact attempts look more like easy-to-reach respondents, while those with more attempts are more similar to difficult-to-reach respondents. Thus, the composition of nonrespondents, as reflected by the distribution of their nonresponse reasons and the number of contact attempts, shapes the overall profile of nonrespondents and, in particular, the extent to which it resembles easier- versus harder-to-reach respondents.

Our findings offer three key practical insights. First and foremost, as we demonstrate in Section 4, nonrespondents' paradata (specifically, their nonresponse reasons and number of contact attempts), combined with data on respondents categorized by difficulty to reach, can help assess nonrespondents' overall profile and, therefore, nonresponse bias.[30]

A second key insight from our findings, driven by the first, is that paradata on both respondents and nonrespondents should be routinely recorded, maintained, and, importantly, made readily accessible to survey users. For instance, in the publicly available Current Population Survey (CPS) data, most nonrespondents are categorized into two broad groups—those who were contacted (Refusals) and those who were not (Absents).[31] In particular, 80 percent of nonrespondents are classified as Refusals, presumably including nonrespondent subgroups such as those with time constraints or other temporary difficulties. Additionally, the count of contact attempts includes only in-person but not telephone attempts, and records a "0'' value for a sixth of respondents and more than a third of nonrespondents. Even in the CE, which offers more detailed paradata, enhancing paradata quality (e.g., regarding the Other and Refused (Other) categories) could improve the potential accuracy of assessing nonresponse bias. Our findings, highlighting the value of high-quality paradata for both respondents and nonrespondents, may motivate national surveys to direct their interviewers to rigorously collect such data, assure its quality, and make it publicly available.

Our final insight is that, counterintuitively, efforts focused solely on increasing response rates to mitigate nonresponse bias may not be as effective as desired, and may even backfire. While surveys often invest resources in boosting response rates (often successfully), our findings indicate that this does not necessarily lead to a reduction in nonresponse bias. We show that nonrespondents in our surveys are a heterogeneous group, with diverse characteristics—including demographics and economic outcomes—

[30]Note that our primary method (applied to LFS and HES outcomes in Section 4.1) requires linked administrative records for outcomes similar to the survey outcomes *aggregated* by paradata (see our Table 4), rather than administrative microdata (which is often inaccessible due to privacy concerns). In fact, aggregated linked administrative data on demographics alone may suffice, as it can help identify similarities between paradata-based groups of nonrespondents and respondents. Furthermore, the method we apply in the CE analysis (Section 4.2) requires only the distribution of nonrespondents by paradata categories (e.g., nonresponse reasons), without relying on linked administrative records at all. However, this method is likely less robust compared to the approach outlined in Section 4.1.

[31]The CPS data dictionary and publicly available datasets are available at U.S. Census Bureau (2023).

that are strongly correlated with (and hence predictable from) paradata such as nonresponse reason. The specific type and intensity of effort employed to increase response rate therefore shape the aggregate composition of nonrespondents, which in turn affects the nature and magnitude of nonresponse bias. For instance, merely increasing contact attempts may disproportionately convert temporarily unavailable nonrespondents, inadvertently moving the resulting profile of respondents *away* from that of the entire population that includes nonrespondents facing more permanent challenges. This insight aligns with several studies challenging the link between nonresponse rate and nonresponse bias, including Curtin et al. (2000); Keeter et al. (2000); Groves (2006); and Küfner et al. (2022). To more effectively reduce nonresponse bias, surveys should consider the specific composition of nonrespondents when deciding where to focus their conversion efforts. Our findings could inform future research aiming to develop specific strategies for doing so.

References

Behaghel, L., Crépon, B., Gurgand, M., and Le Barbanchon, T. (2015). Please call again: Correcting nonresponse bias in treatment effect models. *Review of Economics and Statistics*, 97(5):1070–1080.

Ben-Porath, Y. and Gronau, R. (1985). Jewish mother goes to work: Trends in the labor force participation of women in Israel, 1955-1980. *Journal of Labor Economics*, 3(1, Part 2):S310–S327.

Bollinger, C. R., Hirsch, B. T., Hokayem, C. M., and Ziliak, J. P. (2019). Trouble in the tails? What we know about earnings nonresponse 30 years after Lillard, Smith, and Welch. *Journal of Political Economy*, 127(5):2143–2185.

Curtin, R., Presser, S., and Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly*, 64(4):413–428.

Davern, M. E., Meyer, B. D., and Mittag, N. K. (2019). Creating improved survey data products using linked administrative-survey data. *Journal of Survey Statistics and*

Methodology, 7(3):440–463.

Finkel, Y., Artsev, Y., and Yitzhaki, S. (2006). Inequality measurement and the time structure of household income in Israel. *The Journal of Economic Inequality*, 4(2):153–179.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *International Journal of Public Opinion Quarterly*, 70(5):646–675.

Heffetz, O. and Rabin, M. (2013). Conclusions regarding cross-group differences in happiness depend on difficulty of reaching respondents. *American Economic Review*, 103(7):3001–3021.

Heffetz, O. and Reeves, D. B. (2019). Difficulty of reaching respondents and nonresponse bias: Evidence from large government surveys. *Review of Economics and Statistics*, 101(1):176–191.

ICBS (2018). Household income and expenditure: Data from the 2016 household expenditure survey. `https://www.cbs.gov.il/he/publications/DocLib/2019/1719/e_print.pdf`. Accessed: 12/19/2023.

ICBS (2019). Labour force survey 2017. `https://www.cbs.gov.il/he/publications/DocLib/2019/lfs17_1746/e_print.pdf`. Accessed: 12/19/2023.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., and Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly*, 64(2):125–148.

Kreuter, F., Müller, G., and Trappmann, M. (2010). Nonresponse and measurement error in employment research: making use of administrative data. *Public Opinion Quarterly*, 74(5):880–906.

Kreuter, F., Presser, S., and Tourangeau, R. (2008). Social desirability bias in cati, ivr, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly*, 72(5):847–865.

Küfner, B., Sakshaug, J. W., and Zins, S. (2022). Analysing establishment survey non-response using administrative data and machine learning. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 185(Supplement_2):S310–S342.

Lin, I.-F. and Schaeffer, N. C. (1995). Using survey participants to estimate the impact of nonparticipation. *Public Opinion Quarterly*, 59(2):236–258.

Luiten, A., Hox, J., and de Leeuw, E. (2020). Survey nonresponse trends and fieldwork effort in the 21st century: Results of an international study across countries and surveys. *Journal of Official Statistics*, 36(3):469–487.

Meterko, M., Restuccia, J. D., Stolzmann, K., Mohr, D., Brennan, C., Glasgow, J., and Kaboli, P. (2015). Response rates, nonresponse bias, and data quality: results from a national survey of senior healthcare leaders. *Public Opinion Quarterly*, 79(1):130–144.

Meyer, B. D. and Mittag, N. (2019). Using linked survey and administrative data to better measure income: Implications for poverty, program effectiveness, and holes in the safety net. *American Economic Journal: Applied Economics*, 11(2):176–204.

Meyer, B. D. and Mittag, N. (2021). An empirical total survey error decomposition using data combination. *Journal of Econometrics*, 224(2):286–305.

Meyer, B. D., Mok, W. K., and Sullivan, J. X. (2015). Household surveys in crisis. *Journal of Economic Perspectives*, 29(4):199–226.

OECD (2024). Labour force participation rate (indicator). doi: 10.1787/8a801325-en. Accessed: 6/16/2024.

Rothbaum, J. and Bee, A. (2021). Coronavirus infects surveys, too: Survey nonresponse bias and the coronavirus pandemic. *US Census Bureau*.

Schechtman, E., Yitzhaki, S., and Artsev, Y. (2008). Who does not respond in the household expenditure survey: An exercise in extended Gini regressions. *Journal of Business & Economic Statistics*, 26(3):329–344.

U.S. Bureau of Labor Statistics (2023a). Consumer Expenditure Surveys: Public Use Microdata (PUMD). `https://www.bls.gov/cex/pumd.htm`. Accessed: 12/11/2023.

U.S. Bureau of Labor Statistics (2023b). Household and establishment survey response rates. `https://www.bls.gov/osmr/response-rates/`. Accessed: 12/10/2023.

U.S. Census Bureau (2023). Current Population Survey (CPS): Basic monthly data. `https://www.census.gov/data/datasets/time-series/demo/cps/cps-basic.html`. Accessed: 12/11/2023.

Williams, D. and Brick, J. M. (2018). Trends in US face-to-face household survey nonresponse and level of effort. *Journal of Survey Statistics and Methodology*, 6(2):186–211.

Williams, R. (2011). Using the margins command to estimate and interpret adjusted predictions and marginal effects. In *Stata Conference, Chicago*.

Yashiv, E. and Kasir, N. (2013). Arab women in the Israeli labor market: Characteristics and policy proposals. *Israel Economic Review*, 10(2):1–41.

A Appendix

A.1 Match Quality

This appendix provides a detailed investigation of the difference between dwellings that were matched with administrative records and those that were not. Initially, we delve into the variation in match rates across difficulty-to-reach groups. As illustrated in Table A.1, the match rate monotonically increases from the easiest-to-reach respondents (93.4 percent for the LFS and 88.1 percent for the HES) to the hardest-to-reach (94.7 and 90.9 percent). This trend is primarily attributed to the higher prevalence of Jewish individuals within the hard-to-reach groups, who exhibit a higher matching rate than non-Jewish individuals. In the LFS, the nonrespondents' match rate (93.2 percent) is lower than even the easiest-to-reach respondents, while in the HES, their rate (89.1 percent) is between the first and second contact groups.

Table A.1: Matching to Administrative records by Difficulty-to-reach

Contact attempts		1	2	3+	NR	All
				LFS		
Matched	#	31,411	16,039	19,271	16,917	83,638
	%	93.4	94.2	94.7	93.2	93.8
Not Matched	#	2,222	979	1,079	1,234	5,514
	%	6.6	5.8	5.3	6.8	6.2
Total	#	33,633	17,018	20,350	18,151	89,152
	%	100.0	100.0	100.0	100.0	100.0
				HES		
Matched	#	13,360	9,483	12,641	10,070	45,554
	%	88.1	90.5	90.9	89.1	89.6
Not Matched	#	1,807	998	1,267	1,234	5,306
	%	11.9	9.5	9.1	10.9	10.4
Total	#	15,167	10,481	13,908	11,304	50,860
	%	100.0	100.0	100.0	100.0	100.0

Notes: Samples: Our main LFS and HES samples in Figure 1 and Figure 2, respectively.

Subsequently, Table A.2 employs survey data to present the demographics and outcomes of respondents based on whether they were successfully matched with administrative records ('Matched") or not ("Not Matched"). These two groups are further subdivided based on difficulty-to-reach categories, while the "All" columns provide results encompassing all Matched or Not Matched respondents. See the beginning of

Table A.2: Survey Results—By Matching to Administrative Data

Attempts		Matched 1	Matched 2	Matched 3+	Matched All	Not Matched 1	Not Matched 2	Not Matched 3+	Not Matched All
					LFS survey data				
					A. Demographics				
Age (mean)		44.3	43.9	43.5	44.0	41.7	41.3	41.0	41.4
		(0.1)	(0.1)	(0.1)	(0.0)	(0.2)	(0.4)	(0.4)	(0.2)
Years of Education (mean)		12.8	13.4	13.6	13.2	12.4	13.0	13.6	12.8
		(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)	(0.0)
Female (%)		52.1	52.7	53.1	52.5	51.5	50.6	51.8	51.3
		(0.2)	(0.3)	(0.2)	(0.1)	(0.6)	(1.0)	(1.0)	(0.5)
Jewish (%)		70.7	81.2	87.8	77.8	56.3	64.7	83.3	64.5
		(0.2)	(0.2)	(0.2)	(0.1)	(0.6)	(0.9)	(0.7)	(0.4)
Contacted by Phone (%)		0.2	6.7	44.2	13.7	0.1	6.5	48.8	12.8
		(0.0)	(0.1)	(0.2)	(0.1)	(0.0)	(0.5)	(1.0)	(0.3)
Number of attempts (median)		1	2	4	2	1	2	4	1
					B. Outcomes (Unadjusted Means)				
Labor force participation		58.3	63.6	67.5	62.1	56.1	62.5	67.5	60.2
		(0.2)	(0.2)	(0.2)	(0.1)	(0.6)	(1.0)	(0.9)	(0.5)
Unemployment rate		6.1	5.7	5.4	5.8	8.0	5.6	4.8	6.6
		(0.1)	(0.2)	(0.1)	(0.1)	(0.5)	(0.6)	(0.5)	(0.3)
					C. Outcomes (Adjusted Means)				
Labor Force Participation		61.2	61.6	63.7	61.9	62.0	63.2	62.5	62.4
		(0.4)	(0.2)	(0.2)	(0.1)	(2.4)	(0.8)	(0.9)	(0.4)
Unemployment Rate		5.7	5.9	5.7	5.8	6.3	5.1	5.0	6.1
		(0.4)	(0.2)	(0.2)	(0.1)	(1.7)	(0.6)	(0.7)	(0.3)
					HES survey data				
					D. Demographics				
Age (mean)		49.5	49.1	47.6	48.7	46.2	46.8	44.9	46.0
		(0.2)	(0.2)	(0.1)	(0.1)	(0.4)	(0.5)	(0.5)	(0.3)
Years of Education (mean)		13.2	14.0	14.0	13.7	12.7	13.5	14.1	13.3
		(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.2)	(0.1)	(0.1)
Jewish (%)		76.7	86.4	89.5	83.8	56.1	67.4	77.7	65.6
		(0.4)	(0.4)	(0.3)	(0.2)	(1.2)	(1.5)	(1.2)	(0.7)
Number of attempts (median)		1	2	4	2	1	2	4	2
					E. Outcomes (Unadjusted Means)				
Inc.:	Log NIS	9.37	9.48	9.48	9.44	9.05	9.24	9.29	9.17
		(0.01)	(0.01)	(0.01)	(0.00)	(0.03)	(0.04)	(0.03)	(0.02)
	Converted to NIS	11,732	13,124	13,055	12,558	8,518	10,283	10,784	9,600
		(93)	(125)	(103)	(61)	(233)	(386)	(338)	(174)
Exp.:	Log NIS	9.10	9.17	9.14	9.13	8.99	9.13	9.12	9.06
		(0.01)	(0.01)	(0.01)	(0.00)	(0.02)	(0.02)	(0.02)	(0.01)
	Converted to NIS	8,948	9,597	9,352	9,261	7,988	9,264	9,116	8,631
		(55)	(69)	(59)	(35)	(144)	(209)	(189)	(101)
					F. Outcomes (Adjusted Means)				
Inc.:	Log NIS	9.39	9.43	9.45	9.42	9.36	9.38	9.40	9.38
		(0.01)	(0.01)	(0.01)	(0.00)	(0.02)	(0.03)	(0.02)	(0.01)
	Converted to NIS	11,977	12,440	12,707	12,350	11,604	11,816	12,099	11,796
		(83)	(101)	(94)	(51)	(259)	(322)	(288)	(163)
Exp.:	Log NIS	9.11	9.13	9.13	9.12	9.14	9.16	9.17	9.19
		(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.02)	(0.02)	(0.01)
	Converted to NIS	9,081	9,260	9,214	9,181	9,333	9,537	9,583	9,808
		(42)	(50)	(46)	(25)	(139)	(174)	(152)	(59)
Obs.:	LFS (individuals)	76,945	37,094	42,659	156,698	6,218	2,587	2,656	11,461
	LFS (dwellings)	31,291	15,980	19,128	66,399	2,176	956	1,050	4,182
	HES (dwellings)	13,359	9,483	12,641	35,483	1,807	998	1,267	4,072

Notes: Sample: The LFS Survey Sample (box I in Figure 1) and HES Survey Sample (box IV in Figure 2).

Section 2.1 for a more detailed explanation of the table's structure.

A comparison of the two "All" columns reveals that relative to the Matched group, the Not Matched group is younger (41.4 vs. 44.0 for the LFS and 46.0 vs. 48.7 for the HES), has fewer schooling years (12.8 vs. 13.2 and 13.3 vs. 13.7), and contains a notably lower percent Jewish (64.5 vs. 77.8 and 65.6 vs. 83.8). These patterns also persist when examining each difficulty-to-reach group individually.

Panel C reveals that the Not Matched's adjusted labor force participation and unemployment rate (62.4 [SE 0.4] percent and 6.1 [0.3] percent, respectively) are slightly (and not statistically significant) higher than those of the Matched group (61.9 [0.1] percent and 5.8 [0.1] percent). Additionally, this panel indicates that within the easiest-to-reach group, the Not Matched respondents exhibit higher labor force participation and unemployment rates compared to their Matched counterparts, while this trend reverses for the hardest-to-reach group. For the 2-attempts group, the Not Matched respondents demonstrate higher labor force participation while their unemployment rate is lower.

Panel F shows that the adjusted mean log income of the Not Matched group (11,796 [163] NIS after conversion) is lower than that of the Matched group (12,350 [51] NIS). In contrast, the adjusted mean log expenditure of the Not Matched group (9,808 [59] NIS) surpasses that of the Matched group (9,181 [25] NIS). Similar trends emerge also when scrutinizing each difficulty-to-reach group individually.

These disparities might raise concerns regarding nonmatch bias when utilizing administrative records, similar to nonresponse bias that can arise when using survey data. However, when comparing the data of all respondents (Table 1) with the data from only matched respondents (the Matched section in Table A.2 in the appendix), it becomes apparent that most demographic and outcome differences between these samples are minor. This is mostly attributable to our high matching rate. Consequently, while distinctions between matched and unmatched dwellings are discernible, the impact of excluding unmatched dwellings on our main results seems to be minimal. Nonetheless, it is worth noting that if the nonmatch rate was higher, which might be the case in other survey contexts, nonmatch bias could potentially pose more pronounced repercussions.

A.2 Variations Within Nonresponse Groups

Table A.3: Nonrespondents by Reason (Standard Deviation)

Attempts	Absent	Refused	Temporary	Language	Permanent	Unlocated	Withdraw	Other	All
LFS administrative matched sample									
A. Demographics									
Age	17.7	17.5	16.6	18.2	20.7	21.7		17.6	17.8
Years of Education	3.5	3.5	3.5	5.6	6.2	4.6		3.9	3.7
HES administrative matched sample									
D. Demographics									
Age	18.5	17.5	16.5	18.7	19.8	19.3	18.2	19.2	18.2
Years of Education	3.6	3.4	3.3	5.2	4.6	3.7	3.7	3.3	3.6
E. Outcomes (Unadjusted Means)									
Income: (Log NIS):	2.1	2.2	1.8	2.7	3.8	2.4	2.4	2.0	2.3
F. Outcomes (Adjusted Means)									
Income: (Log NIS):	2.1	2.1	2.1	4.0	3.1	2.2	2.2	2.7	2.2
Obs.: LFS	7,540	3,836	2,998	892	203	399		933	16,917
HES	1,213	3,878	570	215	671	339	3,122	62	10,070

Notes: Sample: the LFS and HES administrative matched samples. Numbers represent standard deviations.

A.3 CE—Quarterly Expenditures by Difficulty to Reach (2021:Q1–2023:Q1)

Table A.4: Respondent Groups by Number of Contact Attempts: 1, 2, 3, 4, 5, 6, 7, 8+

Attempts	1	2	3	4	5	6	7	8+	Total	Total (Wtd.)
Log Expenditures	9.24	9.37	9.47	9.48	9.49	9.53	9.54	9.55	9.46	9.45
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Converted to $	10,311	11,759	12,904	13,065	13,252	13,726	13,967	14,024	12,836	12,758
	(127)	(97)	(104)	(117)	(139)	(162)	(190)	(112)	(44)	(50)
Frequency	3,657	7,666	7,952	6,423	4,841	3,658	2,745	7,757	44,699	44,699

Notes: Source: U.S. Bureau of Labor Statistics (2023a). See details below.

Table A.5: Respondent Groups by Number of Contact Attempts: 1, 2, 3+

Attempts	1	2	3+	Total	Total (Weighted)
Log Expenditures	9.24	9.37	9.50	9.46	9.45
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Converted to $	10,311	11,759	13,415	12,836	12,758
	(127)	(97)	(53)	(44)	(50)
Frequency	3,657	7,666	33,376	44,699	44,699

Notes: Source: U.S. Bureau of Labor Statistics (2023a). See details below.

Table A.6: Respondent Groups by Number of Contact Attempts: 1–2, 3–4, 5+

Attempts	1–2	3–4	5+	Total	Total (Weighted)
Log Expenditures	9.33	9.47	9.53	9.46	9.45
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Converted to $	11,270	12,976	13,758	12,836	12,758
	(78)	(78)	(71)	(44)	(50)
Frequency	11,323	14,375	19,001	44,699	44,699

Notes: Source: U.S. Bureau of Labor Statistics (2023a). See details below.

Source for CE Tables: The source the tables in this appendix section (as well as Table 10) is the U.S. Bureau of Labor Statistics (2023a), which includes publicly available datasets and detailed information. For Appendix A.3, the number of contact attempts is based on the number of times a household appears in the "MCHI" file, which includes data related to the contact history. The expenditures data covers up to 95 percent of the household expenditures over the past three months (calculated as the sum of the “TOTEXPCQ” and “TOTEXPPQ” variables from the “FMLI” file). The weighted average of log expenditure for the entire sample of respondents is computed using the survey weights (variable “FINLWT21” from the “FMLI” file). For Table 10, the reason for nonresponse is determined based on the final interview status (as indicated by the “OUTCOME” variable from the “FBAR” file).

A.4 Extrapolated Expenditures for Nonresponse Categories

Table A.7: Alternatives of Extrapolating Expenditures for Nonresponse Categories

Alter.	Absent	Refused (Hostile)	Refused (Other)	Time Constraint	Language	Other	Average (Wtd.)
1	9.55	9.11	9.24	9.55	9.11	9.11	9.33
2	9.55	9.11	9.24	9.55	9.11	9.11	9.33
3	9.55	9.24	9.24	9.55	9.24	9.24	9.36
4	9.55	9.24	9.24	9.55	9.24	9.24	9.36
5	9.50	9.11	9.24	9.50	9.11	9.11	9.31
6	9.64	9.11	9.24	9.64	9.11	9.11	9.37
7	9.50	9.24	9.24	9.50	9.24	9.24	9.34
8	9.64	9.24	9.24	9.64	9.24	9.24	9.39
9	9.53	9.19	9.33	9.53	9.19	9.19	9.38
10	9.59	9.19	9.33	9.59	9.19	9.19	9.40
11	9.53	9.33	9.33	9.53	9.33	9.33	9.41
12	9.59	9.33	9.33	9.59	9.33	9.33	9.43
Percent	22.3	12.2	39.1	16.5	1.4	8.3	100.0

Notes: This table presents the assumed log expenditures of each nonresponse category under each alternative in Table 11.

A.5 CE—Nonresponse Reason by Attempts

Table A.8: CE: Nonresponse Reason by Attempts - Percent (2021:Q1–2023:Q1)

	Absent	Refused (Hostile)	Refused (Other)	Time Constraint	Language	Other
1	1.7	7.5	4.2	3.5	6.8	34.6
2	5.4	12.5	8.9	6.6	10.6	8.4
3	8.4	13.3	11.5	8.8	10.8	8.3
4	10.4	12.9	12.2	10.2	13.1	7.2
5	11.7	11.1	11.5	10.9	12.5	6.3
6	11.2	9.5	10.3	10.2	9.8	5.6
7	10.4	7.1	8.6	9.0	9.5	5.2
8+	40.8	26.0	32.9	40.7	26.9	24.3
Total	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Source: U.S. Bureau of Labor Statistics (2023a). For nonresponse reason, the variable "OUTCOME" from the "fpar"' file was utilized. The frequency of contact attempts was determined based on the number of times an individual appeared in the "mchi" file.