

**Finance and Economics Discussion Series  
Divisions of Research & Statistics and Monetary Affairs  
Federal Reserve Board, Washington, D.C.**

**An Assessment of the National Establishment Time Series  
(NETS) Database**

**Keith Barnatchez, Leland D. Crane, and Ryan A. Decker**

**2017-110**

Please cite this paper as:

Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker (2017). "An Assessment of the National Establishment Time Series (NETS) Database," Finance and Economics Discussion Series 2017-110. Washington: Board of Governors of the Federal Reserve System, <https://doi.org/10.17016/FEDS.2017.110>.

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

# An Assessment of the National Establishment Time Series (NETS) Database

Keith Barnatchez, Leland D. Crane, and Ryan A. Decker\*

October 27, 2017

## Abstract

The National Establishment Time Series (NETS) is a private sector source of U.S. business microdata. Researchers have used state-specific NETS extracts for many years, but relatively little is known about the accuracy and representativeness of the nationwide NETS sample. We explore the properties of NETS as compared to official U.S. data on business activity: The Census Bureau's County Business Patterns (CBP) and Nonemployer Statistics (NES) and the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW). We find that the NETS universe does not cover the entirety of the Census-based employer and nonemployer universes, but given certain restrictions NETS can be made to mimic official employer datasets with reasonable precision. The largest differences between NETS employer data and official sources are among small establishments, where imputation is prevalent in NETS. The most stringent of our proposed sample restrictions still allows scope that covers about three quarters of U.S. private sector employment. We conclude that NETS microdata can be useful and convenient for studying static business activity in high detail.

## 1 Introduction

We explore the representativeness of the National Establishment Time Series (NETS), a private sector source of business microdata, relative to official U.S. business universe data

---

\*Barnatchez is a student at Colby College. Crane and Decker are economists at the Federal Reserve Board. Bo Yeon Jang provided excellent research assistance. Without implication, we thank Maria Tito for helpful conversations. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the Federal Reserve research staff or the Board of Governors.

sources: The Census Bureau's County Business Patterns (CBP) and Nonemployer Statistics (NES) and the Bureau of Labor Statistics' (BLS) Quarterly Census of Employment and Wages (QCEW). A key purpose of the present note is to provide background for the exercises described in [Crane et al. \(2017\)](#), which uses 2014 NETS data to estimate flood exposure during and after Hurricane Harvey, but our analysis will be useful for potential NETS users more broadly.

NETS consists of establishment-level longitudinal microdata covering, in principle, the universe of U.S. businesses. Though costly, NETS can be accessed without extensive proposal and security clearance processes and can be used outside of secure government facilities, potentially providing a highly efficient way to conduct research on topics that require business-level microdata. However, NETS data are not generated by the rigorous processes that characterize official data collection activities of U.S. statistical agencies, and comparisons of NETS to official sources have raised questions about the specific business universe covered by the data as well as the quality of annual-frequency information on establishment-level entry, exit, and employment (see, e.g., [Haltiwanger et al. \(2013\)](#) and [Neumark et al. \(2005\)](#), which we discuss further below).

A number of studies have documented properties of NETS in limited, single-state samples. Fewer studies have used the full national NETS file, and these studies typically restrict the national file to a subset that can be matched to external data samples. Using national NETS data for 1992-2014, we highlight previous concerns about the precise nature of the NETS business universe, document limitations in terms of imputation and other data artifacts, and outline specific sample restriction criteria that render NETS reasonably comparable to official sources for the purpose of studying static business distributions. In related work in progress not reported here, we find that NETS is much more limited in its value for studying business dynamics.

After applying appropriate sample selection criteria, we find that the correlation of NETS employment counts and CBP employment counts across U.S. counties can be in excess of

0.99. Correlations across state-industry-size class cells are somewhat lower but are still above 0.9 on restricted samples. Zip code-level correlations are also remarkably high. That said, we find several discrepancies between NETS and the other data sources. NETS heavily over-represents establishments with fewer than 10 employees relative to official employer business data, possibly due to the imputation of positive employment to nonemployers in NETS. We also find smaller but significant discrepancies among the largest establishments, likely due in part to large, public educational establishments that are difficult to identify as government owned, as well as establishments that mistakenly report firm employment numbers in NETS. Finally, post-2000 developments in U.S. mining, construction, and manufacturing employment do not appear to be well captured in NETS, though these discrepancies may be due in part to industry labeling differences. When we omit very small establishments, NETS agrees reasonably well with CBP and QCEW, both in terms of trends over time and in terms of the cross section, and we find better agreement still when we omit the largest establishments and educational establishments. This most restrictive sample accounts for 73 percent percent of US employment in the QCEW. We propose that any NETS results be checked on both the unrestricted sample and the restricted sample to ensure robustness.

The paper proceeds as follows. Section 2 describes NETS and the related literature. Section 3 contains the main results, and Section 4 concludes. Additional tables and figures, as well as details regarding CBP and QCEW, are found in the appendices.

## **2 NETS background**

NETS is a product of Walls & Associates. The source data for NETS are collected by Dunn & Bradstreet (D&B) for the Duns Marketing Information file (DMI; see Walls (2008)). D&B uses and sells the data for, among other things, marketing and credit scoring. While there is no legal obligation for establishments to participate or report truthfully, D&B has strong profit-based incentives to compile accurate data, and individual businesses' access to credit

and other business relationships may depend on the quality of the information they provide. D&B attempts to collect information on all U.S. businesses from secretaries of state, Yellow Pages, court records, and credit inquiries, as well as other sources. They also contact businesses directly by telephone. Each establishment is assigned a unique *dunsnumber*, which is constant over time and follows the business when it moves or when it is acquired by another firm. D&B also attempts to link establishments to firms; NETS files include an *hqduns* number for each establishment, which is the dunsnumber of the ultimate domestic headquarters.<sup>1</sup>

While D&B concepts are not as rigidly defined as concepts used by statistical agencies, a reasonable assessment is that D&B attempts to catalog every business establishment in the U.S., where “business” is broadly defined to include private for-profit and nonprofit organizations as well as government agencies. In NETS, an establishment is a specific *line of business* at a specific location (see below), and employment includes all workers at an establishment, potentially including proprietors, independent contractors, and temporary workers supplied by outside organizations. While this broad scope suits D&B business purposes well, it can be difficult to reconcile with well-defined universe concepts employed by U.S. statistical agencies; as we show below, though, NETS data are sufficiently detailed to allow approximation of standard business scope definitions. For example, the NETS line of business concept is a subset of standard establishment concepts in official sources, in which an establishment includes all workers at a specific business location (and, hence, is the sum of all lines of business in operation at that location).<sup>2</sup>

An additional difference between NETS and official sources is that official sources record establishment existence and employment on specified, uniform dates (for example, the pay period including March 12). NETS records are annual, but information is collected throughout the year, and the timing of measurement for each establishment is not reported in the

---

<sup>1</sup>There is no information in our files on the intermediate layers of the firm. *Hqduns* points to the highest-level headquarters within the U.S., never to, e.g., the regional headquarters of the firm.

<sup>2</sup>In our baseline results, we merge NETS lines of business to better approximate the CBP/QCEW establishment concept. All of our qualitative conclusions are unchanged if instead we use the separate lines of business.

data. This is not only a source of micro-level measurement error but also a likely cause of discrepancies between NETS aggregates and official data sources.

Besides firm linkages, D&B collects and constructs numerous establishment characteristics; for our purposes, the most important are employment, industry, and location (address). Notably, establishment employment generally includes the firm owner(s). Employment information usually comes from direct inquiries by D&B, unless it is imputed (employment imputation codes are included in NETS files, and we discuss imputation rates below). Industry is generally either self reported by the business or drawn from Secretaries of State data. Industry is available as 8-digit SIC or 6-digit NAICS codes, though there are no imputation codes for industry information. Address information may be self reported or collected from administrative-type records such as Yellow Pages and includes street address, zip code, state and county, as well as latitude and longitude.

## 2.1 Related Literature

NETS data have been used in a number of studies. A key reference is [Neumark et al. \(2005\)](#), which provides a detailed discussion of the history of D&B data collection, and then compares California NETS data to several official sources. [Neumark et al. \(2005\)](#) appears to be the source of the rules of thumb that (a) the employer universe can be approximated by subtracting 1 from all NETS establishments' employment, and (b) business dynamics are best studied at 3-year frequency instead of annually (in a companion paper, we further explore business dynamics in NETS). A number of subsequent studies adopt these conventions. In their California sample, [Neumark et al. \(2005\)](#) find that total NETS employment is not dramatically different from the sum of UI-based employer establishment count measures and the Census Bureau's nonemployer counts (as we detail below, though, in the national sample we find that total NETS employment is consistently between the size of the employer universe and the union of employer and nonemployer universes, a concern noted by [Haltiwanger et al. \(2013\)](#)). Moreover, the authors report high correlations of employment levels

between NETS and official employer datasets at the county-by-industry level and, to a somewhat lesser extent, the industry-by-size level. Consistent with the notion that NETS includes a large number of nonemployer businesses, the authors show that differences between NETS and employer datasets are heavily concentrated in small establishment size classes, particularly the 1-4 employee class. Importantly, [Neumark et al. \(2005\)](#) find considerable evidence of rounding (to the nearest 10 or 5) in NETS employment numbers; moreover, the authors document a significant amount of employment imputation, particularly in establishments' early years in the data, and they show that annual changes in industry-by-county employment are only weakly correlated between QCEW and NETS (correlation 0.53).<sup>3</sup> A handful of subsequent studies by Neumark and coauthors rely on the California NETS sample (e.g., [Kolko and Neumark \(2010\)](#)). [Neumark et al. \(2011\)](#) study business growth in the national sample of NETS but refer to the previous California-based results for background on representativeness.

Another key reference for understanding representativeness in NETS is [Choi et al. \(2013\)](#) (and the follow-up paper [Choi et al. \(2017\)](#)), which uses the Georgia NETS extract. The authors focus primarily on comparisons with CBP and, importantly, use industry codes, establishment size, and legal status criteria to create a NETS universe that is roughly consistent with CBP scope. We follow a similar approach in the present study, creating a NETS sample that is restricted to match CBP scope as closely as possible. Another significant investigation is [Echeverri-Carroll and Feldman \(2017\)](#), who use secretary of state data for Austin, Texas and the North Carolina Research Triangle to validate NETS founding dates; after extensive efforts using automated and hand matching, the authors find that about half of NETS founding dates match secretary of state founding dates (which reflect business formation applications), with 75 percent of matches being accurate within two years and 80 percent of

---

<sup>3</sup>Three-year changes in industry-by-county employment have a correlation of 0.86. [Neumark et al. \(2005\)](#) further study the business dynamics properties of NETS by obtaining founding date data on two sets of businesses (San Francisco phone listings and BioAbility biotech records) and comparing these external data sources with NETS. The authors find that NETS founding dates are accurate about three quarters of the time and are within two years of accuracy about 90 percent of the time; they do not investigate firm size at birth.

matches being accurate within three years. These results confirm the concerns of [Neumark et al. \(2005\)](#) and [Haltiwanger et al. \(2013\)](#) about recognition of founding dates and annual-frequency dynamics.<sup>4</sup> A particularly important contribution of [Echeverri-Carroll and Feldman \(2017\)](#) is the direct comparison of two different NETS vintages, allowing for a study of revision history; the authors find significant establishment additions between the 2013 and 2014 NETS vintages, with large revisions extending back more than four years from the end of the time series. This suggests the need for caution in interpreting data near the end of a NETS sample (the authors argue for dropping the last two years of coverage).<sup>5</sup>

Exploration of the national NETS sample has been more limited. [Mach and Wolken \(2012\)](#) use NETS along with the Survey of Small Business Finance (SSBF), which uses the DMI as its sampling frame; the authors restrict attention to NETS records matched to the 2004 SSBF sample, and they note that there are some employment discrepancies between NETS and D&B-based SSBF data. [Amezcuca \(2010\)](#) uses the national NETS file but restricts it to records that can be identified as startup incubators (relying in part on external data). [Acs et al. \(2008\)](#) use the DMI files on which NETS is based, constructing their own longitudinal version of the data and studying high-growth firms. In principle this independently constructed longitudinal file should be similar to NETS, though no direct comparison has been undertaken. These authors rely heavily on business dynamics data in the DMI, raising some concerns based on high-frequency limitations of the data found in other studies. [Greenstone and Mas \(2012\)](#) use an extract of the national NETS file, providing some limited comparisons with official sources, though they switch to the LBD in subsequent revisions ([Greenstone et al. \(2014\)](#)).

Aside from the above, we are not aware of any significant attempts to benchmark the national NETS file to official data sources.

---

<sup>4</sup>It is important to note, though, that founding date is a difficult concept more generally given movement between employer and nonemployer universes (see, e.g., [Davis et al. \(2009\)](#)).

<sup>5</sup>Several other studies use state-specific NETS extracts. For example, [Cromwell \(2015\)](#) studies Miami-Dade county living wage contracts using the Florida NETS extract. [Currie et al. \(2010\)](#) and [Groizard et al. \(2015\)](#) use the California extract to study fast food locations and manufacturing establishments, respectively. [Donegan \(2014\)](#) studies biomedical businesses in the North Carolina Research Triangle.



### 3 Methodology and Results

We compare NETS to various official data sources and try to understand, to the extent possible, the sources of discrepancies. We focus on finding ways to align NETS with CBP and QCEW. Error is present in all data sources, but CBP and QCEW have important advantages relative to NETS. These programs are based on consistent, well-documented methodology, with the explicit goal of producing representative data. While the coverage and methods may not always be ideal, it is invaluable to have data generated by a well-understood collection process. In contrast, the source data for NETS are collected as a consequence of D&B's other business processes. As such, the data collection methodology will likely never be as transparent or longitudinally consistent as CBP or QCEW.

That said, NETS is certainly an impressive effort by the private sector to construct a research-ready database, and we will argue that it can complement official data sources with some caveats. NETS provides geographic detail and other features that are not available in the most easily accessible government datasets. To the extent that NETS is shown to agree with official data on common scope concepts, we can feel confident extending analyses using the unique features of NETS. For example, CBP has been used at the county level to estimate storm effects (e.g., [Bayard et al. \(2017\)](#)); these kinds of analyses can be extended to sub-county geographies using NETS, as in [Crane et al. \(2017\)](#). This extension is validated by the fact that NETS and CBP are largely in agreement with respect to the geographic distribution of businesses.

#### 3.1 Official data sources

We use three official sources: County Business Patterns (CBP), Nonemployer Statistics (NES), and the Quarterly Census of Employment and Wages (QCEW). All of these data sources are freely available to the public, and these comprise the main official sources of information about the business universe in the U.S. Note that fine aggregations of public use CBP and QCEW files are occasionally censored to protect confidentiality. In these cases, we impute

employment numbers using the national average for the relevant establishment size cell. This censoring occurs *only* for employment, so establishment counts are unaffected.

As noted above, NETS uses different employment and establishment concepts and timing than official data sources. The most well-known government data sources, including CBP and QCEW, cover nearly all workers who receive a regular paycheck but exclude many business owners, the self employed, and independent contractors. NETS, in principle, includes all of these groups, making its employment and establishment universes proper supersets of CBP and QCEW. In addition, CBP excludes most government employment, and many QCEW tabulations exclude all government employment, whereas NETS includes government. Though there is no explicit ownership code that distinguishes government from private establishments in NETS, most government establishments can be flagged in NETS by using NAICS codes and firm linkages, as we discuss below. Finally, annual data from official sources reflect snapshots at a certain time of the year (March 12 in CBP, a date that can also be observed in the monthly QCEW data). NETS data can be collected at any time of the year, and the timing is not disclosed in the data.

### 3.1.1 County Business Patterns (CBP)

The Census Bureau's CBP program provides (publicly available) annual tabulations of establishment counts, employment counts, and payroll by geography, industry and establishment size class based on mid-March snapshots (i.e., employment information reflect the payroll period including March 12). We focus on post-1997 CBP data, after the program switched from SIC industry classification to NAICS. The source data for CBP is the Census Bureau's Business Register, which is in turn built from federal business tax records, surveys, and the Economic Census (conducted in years ending in 2 and 7).<sup>6</sup> Importantly,

---

<sup>6</sup>The Business Register is also the source for the Longitudinal Business Database (LBD) and the Business Dynamics Statistics (BDS), the workhorse datasets for the study of business dynamics in the U.S. (see [Jarmin and Miranda \(2002\)](#)). Access to LBD microdata (as well as Business Register files) requires an approved research proposal and special sworn status, while the BDS (which consists of various aggregations of LBD data) is publicly available. BDS data are constructed to match the scope of CBP. See [DeSalvo et al. \(2016\)](#) for details on the construction of the Business Register.

the Business Register is based on IRS and Social Security Administration (SSA) lists of all known businesses in the U.S., and employment data are derived from these federal sources.<sup>7</sup> Single-establishment firms are efficiently covered with tax records supplemented at times with industry and location information from survey and census data. Information on multi-establishment firms also comes from tax records supplemented by census and survey data. The Economic Census and the annual Company Organization Survey are particularly critical for tracking multi-unit status and distributing employment across establishments.

CBP covers nearly all non-government employer businesses; the exact coverage exceptions are listed in Section A.1 in the appendix. Employment includes all wage and salary workers, both full- and part-time, and excludes proprietors, partners, independent contractors, and temporary help service workers employed by outside establishments (the latter of which are included in the establishment that issues their paycheck rather than the establishment where they work). In addition to excluding self-employed individuals, the CBP excludes private households, railroads, agricultural production employees, and government employees. A popular misconception is that CBP excludes sole proprietorships; some sole proprietorships have payroll employees and therefore can appear in the CBP. The files we use are limited to the 50 states and the District of Columbia.

### **3.1.2 Nonemployer Statistics (NES)**

The Census Bureau's NES shares the same industry scope as the employer statistics used in the CBP and is thus the nonemployer counterpart to CBP. All entities with taxable business income but no employees comprise the total potential set of nonemployers, but Census procedures remove those nonemployer tax entities that can be connected to multi-unit employer businesses as well as regulated investment companies such as mutual funds. Further removals are based on revenue thresholds that vary by legal form; businesses with

---

<sup>7</sup>The Business Register's reliance on federal tax data for employment information contrasts with the BLS employer universe files, which rely on separate state unemployment insurance data (as we detail below). However, the Business Register is supplemented with some geographic and industry information from state-based records provided via BLS.

less than \$1,000 in revenue are dropped, as are businesses above certain thresholds (\$1 million for non-services corporations and partnerships, \$2 million for services corporations and partnerships, and industry-dependent cutoffs for sole proprietorships).<sup>8</sup> In principle, these thresholds could give rise to discrepancies between Census data and NETS to the extent that the latter captures nonemployers with extremely low or high revenue.

NES is available in county-by-industry aggregations. Note that the union of the CBP and the NES, roughly speaking, comprises the universe of U.S. businesses as known to the IRS, with the exception of businesses in out-of-scope industries.<sup>9</sup>

### 3.1.3 Quarterly Census of Employment and Wages (QCEW)

The QCEW is based on the BLS' independent counterpart to the Census Bureau's Business Register. The BLS data are derived from state unemployment insurance (UI) records supplied by State Workforce Agencies; BLS collects monthly data on employment (collected quarterly but covering the pay period including the 12th of each month) from these state sources. The BLS supplements UI records with frequent surveys of multi-establishment firms as well as a rotating survey on industry and geographic information. Inclusion in the QCEW is based primarily on whether an organization is part of the UI system; this is mandatory for most for-profit businesses but optional for some nonprofits.<sup>10</sup> Employees in QCEW, as in CBP, are wage and salary workers (both production and supervisory, with few exceptions); proprietors and other self-employed individuals (who have no employees), independent contractors, some farm and household workers, and externally supplied temporary workers are not included (though the latter are counted as employees of the agencies

---

<sup>8</sup>Changes in nonemployer screening mechanisms were implemented in 2009 such that there is a modest break in the time series at that time. The new screening mechanisms have not been retroactively applied to previous years with the exception of 2008, in which the changes affected about 0.2 percent of firms.

<sup>9</sup>The source data for the NES is also used to construct the confidential Integrated Longitudinal Business Database (ILBD), in which nonemployer businesses are linked longitudinally and combined with the employer universe longitudinal file, the LBD (see [Davis et al. \(2009\)](#), who describe the ILBD and document considerable movement of businesses between the nonemployer and employer universes).

<sup>10</sup>Note that the UI source data underlying the QCEW are also used to construct the publicly available BLS Business Employment Dynamics data as well as the Census Bureau Longitudinal Employer-Household Dynamics data and their public-use aggregate, the Quarterly Workforce Indicators.

that supply them). QCEW includes some non-employers in establishment counts, since UI accounts that had paid employees in previous quarters are sometimes retained in the state databases after becoming nonemployers. As is the case for CBP, it is a misconception that QCEW excludes all sole proprietorships; those who hire payroll employees are typically subject to the UI system and are therefore likely to appear in QCEW. High-level QCEW tabulations include government employment, though in our work we always restrict to the private sector. We also restrict the QCEW sample to the annual observation covering March 12th to be consistent with CBP timing. As we note above, record timing in NETS is unknown and can vary by establishment and year.

In terms of industry coverage, QCEW is neither a superset nor a subset of CBP. A detailed list of QCEW industry coverage is in Section A.2. [Becker et al. \(2005\)](#) provide further discussion of differences between QCEW and CBP and show that there remain small but nontrivial discrepancies between the two sources (in terms of both establishment and employment counts) even after harmonizing industry scope. An advantage of the U.S. statistical system is the availability of both QCEW and CBP, both of which serve as universe files (subject to minor scope restrictions) yet provide almost entirely independently generated information about U.S. business activity.

Figure 5 in the appendix reports aggregate employment in CBP and QCEW in samples that restrict both data sources to industries in the intersection of their respective industry scopes. These scope restrictions include complete omission of industries that are partially out of scope since determining the exact overlap between the two sources is impossible in such industries. We see in Figure 5 that CBP has higher employment than QCEW, and this does not change markedly over time. Tables 6 and 7 in the appendix show these differences by size class, where it is apparent that employment discrepancies between CBP and QCEW reflect, primarily, measurement of large establishments and, secondarily, measurement of very small establishments. Multi-establishment firms in the QCEW are measured more frequently and with different survey designs than in CBP, and precise measurement

of small establishments is difficult in any source due to the grey area between employer and nonemployer status. In addition, it is known that QCEW includes some zero-payroll establishments, which are entirely excluded from CBP. The differences between CBP and QCEW are fairly sensitive to moderate changes to scope rules. Further investigation of these differences is beyond the scope of this paper; see [Fairman et al. \(2008\)](#) for more detailed discussion. While these differences are significant and warrant further research, we will see that the differences between NETS and either CBP or QCEW are an order of magnitude larger.

### **3.2 Analysis samples**

To address the differences in coverage and definitions outlined above, we construct several analysis samples from the full NETS files.

#### **Unrestricted**

This sample includes all establishments and workers in NETS. Our only modification to the data is to merge certain NETS establishments to make the establishment concept closer to that of CBP and QCEW by locating NETS “lines of business”, which represent different portions of an enterprise that are located in the same location. We identify dunsnumbers (i.e., lines of business) that have the same hqhduns (headquarters identifier), 5-digit zip code, and first five street address characters (this latter item amounts to the building number and the first one or two characters of the street name). The purpose of this filter is to identify dunsnumbers that are in the same building and the same firm. We want to abstract from slight variations in the address (e.g. “ST” vs “STREET”) and differences in suite numbers. Thus, we rely on the zip code and the truncated address to make the match, and the matching firm criterion precludes the spurious matching of independent businesses that operate in the same building. We merge (sum) the employment of appropriately matched lines of business, treating them as a single establishment and assigning the merged establishment the NAICS code of the largest line of business (in terms of employment). This should

roughly correspond with the principle activity industry concept observed in official data.

### **Baseline-CBP and Baseline-QCEW**

Beginning with the unrestricted sample (in which lines of business have already been merged), we first identify and exclude government establishments as those with NAICS 92, those with a headquarters that has NAICS 92, or those in a firm where more than half of establishments have NAICS 92. Again, this is motivated by the fact that CBP and most QCEW statistics exclude government establishments and employment. We then restrict the industry scope of NETS to match each respective official data source, resulting in -CBP and -QCEW variants of the baseline sample (the coverage details are in Section A of the appendix).

In addition, it is worth noting that several industries are *partially* covered by QCEW. The extent of undercoverage is not documented in detail, so we completely exclude these industries from both QCEW and baseline-QCEW NETS sample. See Section A of the appendix for details.

### **Unmerged Lines of Business (Unmerged LoB-CBP and Unmerged LoB-QCEW)**

These samples are identical to the baseline samples (including CBP or QCEW industry scope restrictions) but retain all NETS lines of business as separate establishments. As with the baseline samples, we construct two variants that mimic the coverage of CBP and QCEW respectively.

### **Excluding NAICS 61 (Ex. 61-CBP and Ex. 61-QCEW)**

As we will discuss, it appears that many large public universities and school systems that are outside the scope of official data are not captured by our government filter. It is difficult to distinguish public educational institutions from private ones, so we instead drop all NAICS 61 establishments from the baseline samples and make comparisons against similarly restricted QCEW and CBP samples. As with the baseline samples, there are two variants that

mimic the coverage of CBP and QCEW respectively.

### 3.3 Treatment of employment

As has been noted, NETS appears to include business owners in their employment counts. [Neumark et al. \(2005\)](#) propose subtracting 1 from the employment of each establishment to align NETS employment with the standard wage-and-salary employment concept. When the establishment only reports a single worker, it is presumed to be a nonemployer and should not be counted in the employer universe. Thus we can estimate the count of employer establishments by dropping all NETS establishments reporting a single worker. This is not a perfect rule, as there likely exist many true employer businesses with only one employee; however, as we show below, the significant portion of the difference between NETS and official sources is found among the smallest establishments, and official sources indicate that the nonemployer universe is large compared to the small-employer universe. As such we believe that identifying nonemployers in this way implies less measurement error than treating them as employers.

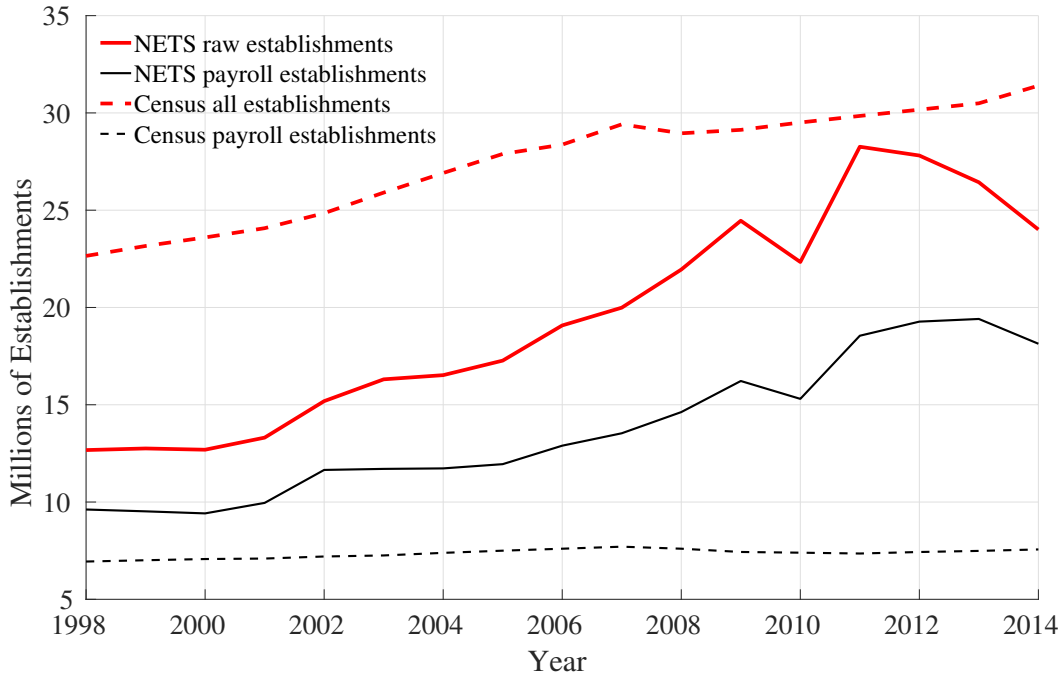
We modify the [Neumark et al. \(2005\)](#) rule slightly, subtracting 1 from the employment of each *headquarters* establishment rather than all establishments. This is motivated by the presumption that the non-payroll owner would work only at the firm headquarters. In practice, the outcomes of using our rule are nearly indistinguishable from those using the broader [Neumark et al. \(2005\)](#) rule since most firms have only one establishment.<sup>11</sup>

This adjustment can be implemented in any of the samples listed above, so for each of the four samples we obtain two establishment counts and two employment counts. We refer to the *adjusted* numbers as **payroll employment** and **payroll establishments** since they approximate the payroll worker concept of official sources. The raw NETS counts are referred to as **raw employment** and **raw establishments**.

---

<sup>11</sup>It is also possible for firms to have multiple owners, though the prevalence of sole proprietorships suggests that further refinements of our methodology would not substantially change the results.





Source: NETS database, County Business Patterns, Census Nonemployer Statistics.  
 Note: NETS sample restricted to CBP industry scope.

Figure 1: Aggregate establishment counts

### 3.4 Aggregate activity in NETS and Census data

We first focus on comparing NETS with Census data sources. The advantage of focusing first on Census sources is that Census provides data on both the employer universe (CBP) and the nonemployer universe (NES), the union of which is (in principle) the universe targeted by D&B.

Figure 1 plots establishment counts for NETS and for Census data, excluding government. The NETS sample is the baseline-CBP sample, with the thick red line showing raw establishment counts and the thin black line showing payroll establishment counts. “Census payroll establishments” (dashed, thin black line) reflects the payroll employer universe of CBP, and “Census all establishments” (dashed, thick red line) reflects the union of the CBP payroll employer universe and the NES nonemployer universe. In principle, the latter union is conceptually equivalent to the NETS universe.

Comparing the dashed lines, it is evident that the number of nonemployer establishments dwarfs the number of employers in official statistics. The count of raw NETS establishments falls somewhere between the two Census totals, trending closer to the total Census universe in recent years. This suggests that NETS covers more than the employer universe but fails to cover the union of the employer and nonemployer universes, a point noted by [Haltiwanger et al. \(2013\)](#).

In the early years of the sample, the total NETS universe was somewhat too high to match the Census employer universe but far too low to match the total Census business universe. Notably, though, NETS coverage appears to have expanded in recent years. It is unclear whether this reflects changes in targeted scope or truly improved coverage, though we can likely rule out the possibility that the actual U.S. business universe expanded at the rate indicated by NETS over that period. [Figure 1](#) should caution researchers against interpreting rising NETS establishment counts since 2000 as reflecting a surge in business entry.

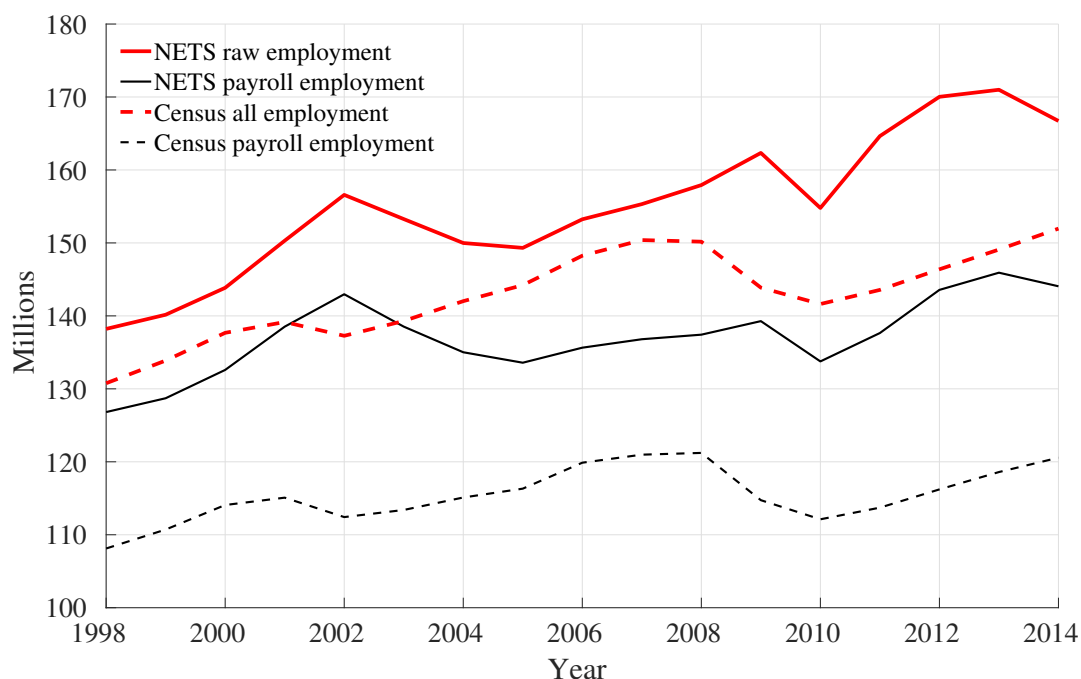
It is also puzzling that NETS counts many more payroll employer establishments (thin, solid black line) than CBP. CBP records are based on IRS and SSA tax data, so there is limited room for mismeasurement. The recent gap implies that about 8 million establishments reported positive employment to D&B but had no wage or salary employees for IRS purposes. We will return to this issue in our discussion of imputation.

[Figure 2](#) plots the employment levels corresponding to the establishment counts in [Figure 1](#). The “Census payroll employment” line is total CBP employment, and the “Census all employment” number sums payroll employment from CBP with the number of business owners in CBP and NES, under the assumption that there is one (non-payroll) owner per establishment.<sup>12</sup> Consistent with [Figure 1](#), NETS payroll employment is significantly higher than Census payroll employment. The raw employment numbers (thick red lines)

---

<sup>12</sup>This differs slightly from our owner adjustment in the NETS data, where we subtract one owner from each *firm*. With CBP and NES alone we cannot implement this adjustment because there are no firm data. In principle we could supplement CBP with firm counts from the BDS, but the difference between the adjustments is very small since most firms only operate a single establishment.

are somewhat closer, but the NETS values are still higher. It is not entirely clear why raw NETS establishment counts are lower than total Census universe establishment counts (see Figure 1), while the corresponding NETS employment counts are *above* the Census counterparts. It appears that while NETS does not cover all the establishments that Census does, NETS must be measuring or imputing higher employment to those establishments it does cover.



Source: NETS database, County Business Patterns, Census Nonemployer Statistics.  
 Note: NETS sample restricted to CBP scope.

Figure 2: Aggregate employment

To summarize, NETS establishment counts fall between the Census employer count and the Census count of all business establishments. The NETS counts begin near the Census employer level in the 1990s but are now closer to the Census employer-nonemployer union. In terms of employment, NETS is consistently above Census, though the raw employment totals are not dramatically different. Taken together, these findings imply that the average *payroll* establishment size in NETS is smaller than in Census. We will turn to establishment

size in the next section.

### 3.4.1 Discussion

There are many possible explanations for these patterns. In the appendix, Figures 7 and 8 show that adding government establishments back into NETS and treating lines of business as separate establishments make almost no difference for the establishment count discrepancies, and Figures 9 and 10 show that the patterns for employment are similarly robust. Another possibility is that the difference between NETS and CBP payroll establishment counts reflects nonemployer establishments that make use of informal workers (such as family), independent contractors, or externally supplied temporary workers; however, this would not explain why total (employer and nonemployer) NETS establishment counts are below the Census business universe. Businesses making use of informal workers may be reluctant to report such to D&B, suggesting that informality is not likely to be quantitatively significant. Instead, below we document high rates of employment imputation among small establishments in NETS, indicating that D&B are not actually receiving employment data from many small businesses. It is therefore unlikely that NETS provides a more accurate count of formal and informal employment.

The problem of independent contractors and temporary workers may be more salient, however, as such workers might be double counted in NETS (i.e., they may be counted both in the establishment in which they work and the establishment that pays them). In unreported results, we construct NETS versus CBP comparisons in which we omit NAICS 56 (which includes temporary help services as well as other commonly contracted business services such as landscaping, janitorial, and security services); the omission of this sector does not materially affect the gap between NETS and CBP.

In what follows, we focus on trying to match NETS payroll employment and payroll establishment totals to official data, rather than raw employment and raw establishment counts, since it is apparent from Figure 1 that NETS does not cover the employer-nonemployer

union. That is not surprising given that D&B does not have access to comprehensive administrative data; larger establishments are likely easier for D&B to locate and perhaps more willing to answer their questions. Therefore, we choose to focus on matching the payroll statistics, dropping presumed nonemployers from the analysis.

We also considered the possibility that the Census nonemployer figures are incorrect. Measuring nonemployers can be difficult, and one has to distinguish between “true” nonemployers and tax entities that may be part of larger firms or financial vehicles.<sup>13</sup> It is conceivable that, for some definition of nonemployer firms, Census counts too many and the NETS figures are closer to truth. We are not aware of any corroborating evidence for this hypothesis, and with the data in hand we cannot test it; however, we note that Census does make serious attempts to restrict the nonemployer universe with linking attempts and revenue cutoffs. Moreover, if Census mismeasurement explains the gap, it must be time-varying mismeasurement since the gap has narrowed. Regardless, the nonemployer universe is a difficult set of businesses to understand, so we feel comfortable restricting attention to payroll employers.

### 3.5 Size distributions in NETS and CBP

To get a better sense of where NETS differs from official data, Table 1 compares payroll employment and payroll establishment counts by size class for 2000, 2007, and 2014. For each year, we report the the percent difference between NETS counts and CBP counts as a percent of CBP counts; for example, the top left cell shows that in 2000, NETS had 82 percent more employment than CBP among establishments with fewer than 5 employees.<sup>14</sup> Several patterns stand out. First, consistent with the discussion in the previous section, NETS employer activity has generally grown over time compared with CBP. The “Aggregate” row shows that in 2000 NETS had 16 percent more employment and 33 percent more establish-

---

<sup>13</sup>See <https://www.census.gov/programs-surveys/nonemployer-statistics/technical-documentation/methodology.html> for the methodology Census uses.

<sup>14</sup>Table 8 in the appendix reports similar exercises but for QCEW instead of CBP.

Size class (Employees)	Percent difference					
	2000		2007		2014	
	Emp.	Estab.	Emp.	Estab.	Emp.	Estab.
1 to 4	82.20	53.85	113.28	130.21	195.87	241.87
5 to 9	15.89	15.24	19.47	18.30	44.21	41.04
10 to 19	4.68	2.94	5.80	4.15	1.29	0.43
20 to 49	8.23	5.93	8.82	6.11	1.85	-0.53
50 to 99	9.98	9.38	8.36	8.07	11.59	9.77
100 to 249	-0.50	0.66	-2.29	-0.27	2.59	4.45
250 to 499	-0.21	0.21	-0.97	-0.07	-3.75	-2.60
500 to 999	9.16	9.30	-1.23	0.42	-5.90	-4.88
1000+	49.36	39.18	13.01	9.80	18.12	3.69
Aggregate	16.23	33.18	13.08	75.66	19.49	139.79
Aggregate ex. less than 10 employees	11.93	4.61	5.10	4.87	4.75	1.33
Aggregate ex. less than 10 and greater than 1000 employees	5.03	4.48	3.63	4.86	2.20	1.32

Source: NETS and CBP

Note: Difference between NETS and CBP employment as percent of CBP employment. NETS sample restricted to CBP scope.

Table 1: NETS versus CBP by establishment size

ments than CBP. By 2014, NETS had 20 percent more employment and more than twice as many establishments as CBP.

The table shows that these differences are primarily driven by small establishments. In 2014 NETS had more than three times as many 1-to-4-employee establishments as CBP, with almost three times as many employees.<sup>15</sup> In the same year NETS had 40 percent more activity in the next establishment size up, those with 5 to 9 employees. More generally, small establishments account for the bulk of the difference between NETS and CBP, and these small establishments also appear to account for much of the expansion of the discrepancy since 2000. This can be seen on the row marked “Aggregate ex. less than 10 employees”, which reports the comparison excluding establishments with fewer than 10 employees from both datasets. In this sample, the differences between NETS and CBP are minor and have

<sup>15</sup>Note again that we have already subtracted one employee from every NETS firm and dropped establishments whose resulting employment is zero.

Size class	Imputation rates		
	2000	2007	2014
1 to 4	41.66	54.85	71.88
5 to 9	21.56	20.17	37.82
10 to 19	19.46	16.82	16.98
20 to 49	17.94	13.83	9.19
50 to 99	16.82	13.75	7.37
100 to 249	14.70	12.89	7.19
250 to 499	18.65	14.91	9.84
500 to 999	18.48	17.77	11.61
1000+	23.74	21.68	15.67

Source: NETS

Note: Establishments with imputed NETS employment as a percent of total NETS establishments by establishment size. NETS sample is restricted to CBP scope but does not merge lines of business.

Table 2: NETS imputation rates by establishment size

not worsened over time. The largest class of establishments, those with 1,000 or more employees, also exhibit some discrepancy. The final row of the table shows that omitting these largest establishments in addition to the smallest ones improves the match modestly, particularly in 2000.

There are a number of possible explanations for the wide disparity among small establishments. We believe that our construction of NETS payroll establishment and employment counts depends on reasonable assumptions about proprietors, but it does introduce minor errors in cases of multiple working (but non-payroll) owners, cases of absentee owners, and cases of on-the-payroll owners (i.e., paycheck-receiving owners who should not be subtracted from employment counts). These errors are less problematic for employment-based comparisons than for establishment count-based comparisons, however, and the divergence from official sources is evident even in the employment numbers. That said, measuring very small businesses (particularly very new businesses, which tend to be small) is a difficult challenge even for official statistical agencies, since there is a fair amount of movement between the employer and nonemployer universes (Davis et al. (2009)) and small businesses

may be more likely to have periods of inactivity.<sup>16</sup> Excluding the small size classes results in a match between NETS and CBP that is not substantially worse than the match between CBP and QCEW (shown on Tables 6 and 7 in the appendix), and the same holds for the match between NETS and QCEW (which we report on Table 8 in the appendix).

The most likely explanation for the divergence among small establishments is the prevalence of imputation in these size classes. Direct contact with the business is an important source of D&B's employment data for these smallest establishments. If the business cannot be contacted or does not answer questions, D&B can be forced to impute missing employment values using cross-sectional information (e.g., establishment location or industry). Walls & Associates reviews these imputations and adjusts them where longitudinal establishment linkages provide information missing from the cross-sectional imputation. Table 2 presents employment imputation rates.<sup>17</sup> Employment is often imputed for the establishments in the smallest size classes (and, of course, size class itself is a function of the potentially imputed employment count).<sup>18</sup> In 2014, more than two-thirds of the employment values for the smallest size class are imputed, and more than one-third of the values for the 5-9 employee class are imputed. Imputation rates in the smallest size class have risen dramatically since 2000, the same period over which excess NETS employment and establishment counts (versus CBP) have risen. The obvious conjecture is that nonemployers are being imputed with positive payroll employment (NETS employment greater than 1), which causes us to treat them as employers. Under these conditions, the simple fix of subtracting 1 from firm employment is not sufficient for satisfactorily reconciling NETS with official employer numbers. Researchers should be cognizant of large measurement error among very small establishments generally, particularly when studying the post-2000 rise of small-establishment activity in NETS.

---

<sup>16</sup>As noted above and shown on Tables 6 and 7, nontrivial discrepancies exist even between CBP and QCEW in the smallest size classes.

<sup>17</sup>These imputation rates necessarily refer to the NETS sample in which lines of business are not merged.

<sup>18</sup>In related work in progress, we find that a significant number of establishments see multiple consecutive years of imputation.



Industry	Percent difference					
	2000		2007		2014	
	Excl <10	All	Excl <10	All	Excl <10	All
11 Ag., For., Fish., Hunt	74	68	67	66	71	73
21 Mining	70	69	-2	3	-8	-5
22 Utilities	-40	-37	-39	-35	-46	-40
23 Construction	-8	3	-5	3	7	23
31-33 Manufacturing	32	34	39	43	50	54
42 Wholesale Trade	7	15	7	17	-4	8
44-45 Retail Trade	-11	-2	-11	-1	-3	4
48-49 Trans., Warehous.	15	19	-6	0	-19	-7
51 Information	15	20	15	21	10	18
52 Finance, Insurance	25	23	5	10	7	12
53 Real Est., Rent., Leas.	70	70	66	68	75	86
54 Prof., Sci., Tech. Svcs	35	39	12	17	5	17
55 Management	-91	-90	-91	-89	-86	-83
56 Admin., Waste Mgmt	-34	-26	-39	-18	-47	-8
61 Education Svcs	287	278	252	244	261	260
62 Health, Social Asst.	5	7	-8	-2	-9	1
71 Arts, Entertain., Rec.	33	44	15	27	2	22
72 Accom., Food Svcs	-1	2	-8	-1	-14	-2
81 Other Svcs	39	41	10	25	6	29

Source: NETS and CBP

Note: Difference between NETS and CBP employment as percent of CBP employment by NAICS sector. NETS sample restricted to CBP scope.

Table 3: NETS versus CBP by sector

### 3.6 Industry distributions in NETS and CBP

To shed further light on differences between NETS and CBP, Table 3 reports employment comparisons by 2-digit NAICS sector. As in our size-based comparisons, we report the difference between NETS and CBP as a percent of CBP levels. For each of 2000, 2007, and 2014, we report the differences both for the full baseline samples (“All”) and for the sample excluding establishments with fewer than 10 workers (“Excl <10”). Within industries, omitting small establishments improves the match in about half of cases.

Educational services (NAICS 61) is the worst-fitting sector, with NETS consistently reporting more than three times as many employees as CBP. This discrepancy has changed

little over time and is likely due to our difficulty identifying large government-owned educational institutions. In QCEW-based comparisons below we find that omitting educational services results in dramatic improvements in aggregate comparisons.

The next significant (though less egregious) discrepancy exists in management of companies and enterprises (NAICS 55); according to the NETS documentation, D&B attempts to avoid using parts of this industry category and instead locates establishments in specific industry areas, resulting in much lower (by more than 80 percent) NETS employment in this sector. Other significant discrepancies exist in agriculture, forestry, fishing and hunting (NAICS 11), in which NETS shows higher employment even after our attempts to limit the sample to CBP scope (which omits much of NAICS 11); this discrepancy has varied only modestly over time.

NETS appears to poorly capture three noteworthy sectoral reallocations of the last 15 years. First, the shale oil and gas boom drove dramatic gains in U.S. mining (NAICS 21) activity starting in the mid-2000s. NETS significantly over-covers this sector in 2000 (by about 70 percent) then approximately matches it in 2007 (the year often chosen to mark the ramp-up of the shale boom) and 2014. While NETS mining coverage may appear to improve over the relevant years, it is far from clear that this coverage expansion actually reflects improved measurement since NETS coverage was so dramatically overstated at the beginning of the sample (and, importantly, this initial overstatement was not driven by small establishments).<sup>19</sup>

Second, NETS appears to miss much of the post-2007 construction “bust” that accompa-

---

<sup>19</sup>Oil and gas industries may be particularly sensitive to industry classification errors. Spot-checking exercises reveal that a number of oil and gas establishments are classified in media-related sectors, perhaps due to the word “production” appearing in their business names; we thank Maria Tito for sharing this discovery with us. This is likely a nontrivial problem but unlikely to be the main driver of the pattern of mining discrepancies. Additionally, since NAICS broad (2-digit) sector classifications often divide similar activities in ways that are effectively arbitrary relative to the set of activities engaged in by businesses, in unreported exercises we created an ad hoc “oil and gas” sector consisting of several relevant narrow industries in mining, manufacturing, transportation, professional services, and construction (designed to encompass oil and gas exploration, production, transportation, refining, and other processing) and compared NETS with QCEW in this specially designed sector. We find that the discrepancies observed in NAICS mining are not significantly altered by the development of this “oil and gas” sector; that is, NETS does appear to truly miss the shale oil and gas boom.

nied the housing crisis (NAICS 23). NETS coverage of the sector is reasonably good in 2000 and 2007, but by 2014 NETS overstates construction employment by just over 20 percent. This overstatement is largely accounted for by small establishments. Weak housing-related coverage may also be seen in finance and insurance (NAICS 52) and real estate and rental and leasing (NAICS 53), in which cases small establishments do not appear to be the main culprit.

Third, NETS seemingly misses some portion of the post-2000 drop in U.S. manufacturing employment (NAICS 31-33). NETS overstatement of the sector rises from about 35 percent in 2000 to about 40 percent in 2007 then to at least 50 percent in 2014. Small establishments account for only a modest portion of the overall overstatement and roughly none of the change over time.

It is important to note that industry classification is a notoriously difficult endeavor. Even in official sources, industry assignment is much less objective than geographic assignment (and even size assignment). As such, some differences in industry coverage between NETS and official sources may partly reflect subjective differences in industry assignment methods rather than substantive measurement error in NETS. While we are aware of little research on this topic, a notable exception is [Isenberg et al. \(2013\)](#). They compare Census Business Register establishment industry codes (i.e., the CBP source data) to the industry in which employees claim to work in the American Communities Survey (ACS). Using links between the worker identifier and the establishment identifier, they find that the worker-reported industry matches the establishment-reported industry 75 percent of the time (at the 2-digit NAICS level). These results suggest that we should expect significant variation in industry assignment across data sources, even at the broadest levels of industry aggregation. The previous section shows that NETS and CBP are reasonably aligned when the smallest size classes are omitted; the industry discrepancies described in this section may therefore partly reflect offsetting differences in subjective industry assignment. That said, the specific discrepancies mentioned here, particularly in mining, construction, and manu-

facturing, are cause for concern (while the differences in education likely reflect true scope differences).

### 3.7 Aggregate activity in NETS and QCEW

In this subsection we compare NETS time series to official data and show how excluding certain classes of establishments can improve agreement between the two. We began this section by focusing on comparisons between NETS and Census data primarily because of the availability of Census nonemployer data, which facilitate full-universe comparison. We now turn to QCEW comparisons. A key advantage of QCEW is that it is available in NAICS format beginning in 1992, so we can easily compare longer time series than with CBP. Moreover, we find that NETS tends to align with QCEW slightly better than with CBP. Recall that we create separate NETS samples meant to mimic the scope of CBP and QCEW, respectively. The better alignment of the NETS baseline-QCEW sample with QCEW suggests that the QCEW-covered industries may be covered more accurately in NETS than the CBP-covered industries. In any case, using CBP instead of QCEW leads to the same qualitative conclusions, and CBP-based versions of the figures in this subsection can be found in the Appendix.

For these comparisons, we restrict attention to employer establishments (i.e., we have already subtracted 1 from all NETS firms' employment and dropped resulting nonemployers) because there is no nonemployer counterpart in QCEW. We focus on the baseline-QCEW sample of NETS that is consistent with QCEW industry scope.

Figure 3 shows time series comparisons of QCEW and NETS in terms of the total number of (employer) establishments and employees. The top row of Figure 3 shows the totals. Consistent with our CBP-based comparisons, NETS reports significantly more establishments and workers than QCEW and is therefore too large to represent the employer universe. Again, the gap between NETS and official establishment counts grows over time. Table 8 in the appendix shows that, as is the case with CBP, the key source of the discrepancy be-

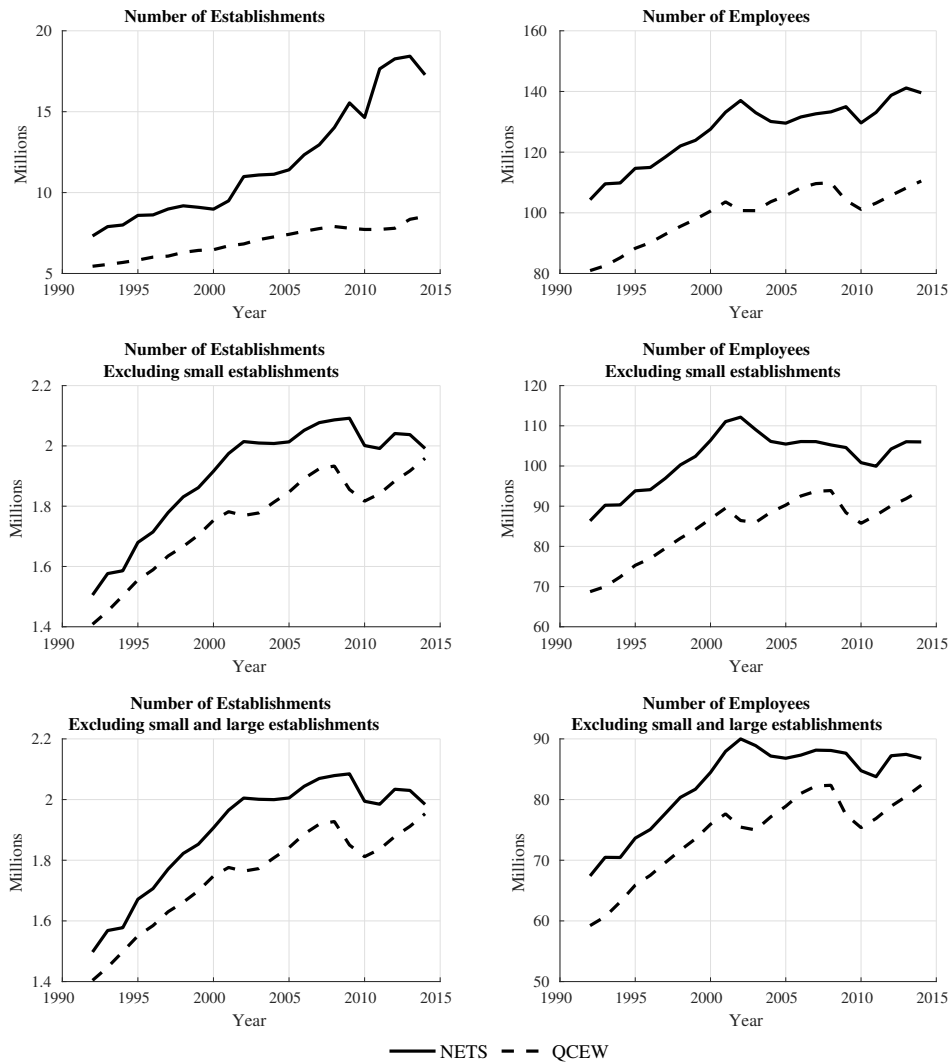
tween NETS and QCEW is small establishments, with additional significant discrepancies among very large establishments. Therefore, the second row of Figure 3 restricts the data to establishments with more than 9 employees, and the third row restricts the data to establishments with more than 9 but fewer than 1000 employees. These restrictions reduce the discrepancy considerably; in particular, small establishments account for the dramatic widening of the discrepancy in recent years. However, under these restrictions NETS still counts more employees and more establishments than QCEW.

Table 9 in the appendix reports differences between NETS and QCEW by NAICS sector; similarly to Table 3 (which compares NETS with CBP), we find significant variation across industries and over time in terms of the discrepancy between NETS and QCEW. As noted with the CBP sector comparisons, a noteworthy sector is educational services (NAICS 61). This sector includes public and private schools as well as martial arts academies and related establishments. Many primary and secondary schools and public universities are not classified as government establishments according to our rule from the initial cleaning process. Thus many large educational establishments are included in our QCEW-comparable NETS sample while they would be excluded from the QCEW private employment figures. Figure 4 repeats Figure 3 but omits education services establishments (in both NETS and QCEW). As the bottom two rows of the figure show, large establishments in the education sector account for a significant portion of the discrepancy between NETS and QCEW, and when we omit both from NETS we are able to replicate QCEW establishment and employment counts reasonably well.

Figures 11 and 12 in the appendix replicate Figures 3 and 4 but for CBP data (with CBP scope), with similar results across the years that are common to both QCEW and CBP.

### **3.8 Comparing narrow cells**

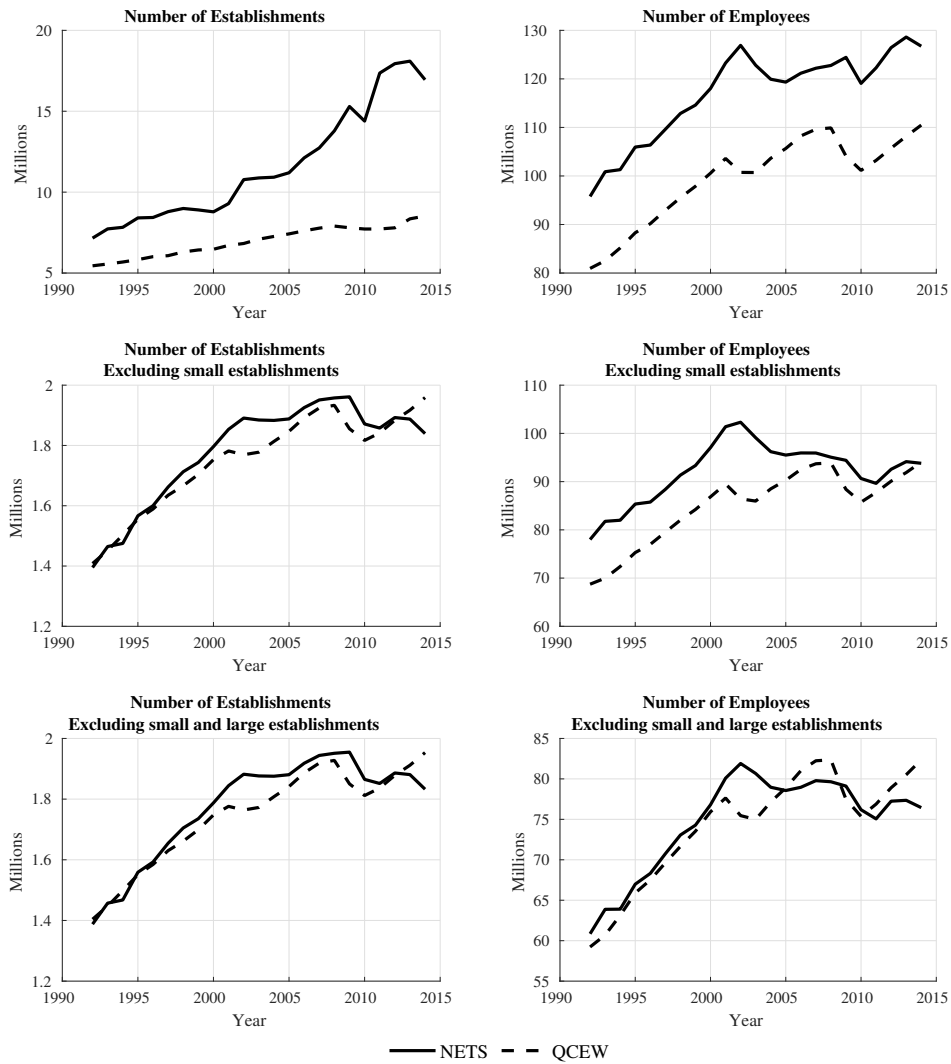
It is not enough for NETS, or a subset of NETS, to have the same employment levels and establishment counts as official sources. To have confidence that inferences from one source



Source: NETS, QCEW.

Notes: NETS sample is restricted to QCEW scope (sample Baseline-QCEW). "Small" establishments are those with fewer than 10 employees. "Large" establishments are those with 1000 or more employees.

Figure 3: NETS versus QCEW



Source: NETS, QCEW.

Notes: NETS sample is restricted to QCEW scope (sample Ex. 61-QCEW). "Small" establishments are those with fewer than 10 employees. "Large" establishments are those with 1000 or more employees.

Figure 4: NETS versus QCEW, education services (NAICS 61) excluded

will carry over to the other, we should have more detailed agreement. To this end, we partition the NETS establishments into various cells and calculate the correlation of cell-level employment and establishment counts with official data. Table 4 presents the results for 2000, 2007, and 2014. The first column notes the official dataset being compared with NETS, the second column notes the level of cell aggregation, and the third column notes sample restrictions (in each case, NETS is restricted to match the scope of the official data to which it is being compared). Remaining columns report simple correlation coefficients.<sup>20</sup> Empty correlation cells on the table indicate that the specified official dataset does not permit suitably accurate correlations at that level due to confidentiality censoring.<sup>21</sup>

The first and fourth rows of the table, corresponding to state-size-sector cells with all size classes in QCEW and CBP respectively, illustrate the problems of mapping the full NETS sample to official sources. Some of these correlations are around 0.5 or 0.6, even though our geography (state) is quite broad. As seen in previous exercises, the relationship between NETS and official sources weakens over time when all sizes are included. The second and fifth rows show, however, that exclusion of small establishments (fewer than 10 workers) both dramatically improves the correlation and significantly attenuates time variation in the match quality. Simply omitting small establishments boosts the correlation between NETS and official sources above 95 percent for establishment counts and above 80 percent for employment. Dropping large establishments—rows three and six—boosts employment correlations to around 90 percent.

<sup>20</sup>For example, let  $n_i^{NETS}$  be the number of establishments that NETS counts in county-size-sector cell  $i$ , and let  $n_i^{CBP}$  be the CBP count of establishments in cell  $i$ . We report the correlation coefficient

$$\rho = \frac{\sum_i \left( n_i^{NETS} - \overline{n^{NETS}} \right) \left( n_i^{CBP} - \overline{n^{CBP}} \right)}{\left[ \sum_i \left( n_i^{NETS} - \overline{n^{NETS}} \right)^2 \times \sum_i \left( n_i^{CBP} - \overline{n^{CBP}} \right)^2 \right]^{1/2}}$$

where the summations and averages are taken over all county-size-sector cells  $i$ , and horizontal bars indicate averages. Employment correlations replace cell establishment counts with cell employment counts.

<sup>21</sup>CBP does not report employment by size class at the county or zip code level, instead reporting employment for the geography-industry cell as a whole. Thus we do not calculate employment correlations for size-by-county or zip comparisons.



The seventh row narrows the comparison to county-based aggregates; county-level employment is nearly perfectly correlated between NETS and CBP. To our knowledge, D&B does not attempt to match county aggregates, so these correlations are strikingly high.<sup>22</sup> This is likely due in part to the fact that business activity is highly correlated with population, so any distinct attempts to measure business activity should be highly correlated across geography; but these results do lend credibility to NETS. As rows 8-10 show, however, the high county-level correlations apparently reflect offsetting misses in narrower (sector-by-size) cells, where correlation is somewhat lower (at this level of detail, only establishment counts are available). As with the state-based correlations, the county-size-sector correlations degrade over time when all size classes are included. However, when small establishments are omitted, the correlations are above 95 percent. The last three rows of Table 4 exploit zip code-level cells from CBP. Cells omitting small establishments see respectable, but not overwhelming, zip-size-sector correlations above 80 percent.<sup>23</sup>

The fact that county aggregates are more correlated than county-size-sector aggregates is noteworthy. At first glance, offsetting misses across size-by-industry cells within counties may be cause for concern. However, as noted in previous exercises, industry is difficult to measure consistently. Unlike geography, industry categorization is necessarily subjective, and caution is always warranted when using industry codes in *any* microdata. High correlation across geography with somewhat lower correlation across industry likely reflects in part (perhaps large part) differences in industry assignment processes. Statistical agencies may be more consistent and rule based when assigning industry codes to individual businesses than is D&B; yet even the best industry measurement processes are subject to considerable ambiguity and error, and D&B does at least have profit incentives to identify businesses in an accurate, useful way. Size measurement is certainly less ambiguous than

---

<sup>22</sup>We confirmed this in correspondence with Don Walls of Walls & Associates. Similar correlations have been noted by Neumark et al. (2005) and others.

<sup>23</sup> These correlations are likely lower bounds since we have 5-digit zip codes from NETS while CBP provides tabulations by zip code tabulation areas (ZCTAs). ZCTAs are often the same as zip codes, but sometimes smaller zip codes are included under larger ZCTA identifiers. This mismatch in identifiers likely makes the measured correlations lower.

Comparison	Cell	Exclusions	Correlation					
			2000		2007		2014	
			Emp.	Estab.	Emp.	Estab.	Emp.	Estab.
QCEW	State-Size-Sector	None	0.82	0.79	0.88	0.61	0.81	0.49
QCEW	State-Size-Sector	<10	0.81	0.97	0.89	0.97	0.86	0.96
QCEW	State-Size-Sector	<10, >1000	0.91	0.97	0.91	0.97	0.87	0.96
CBP	State-Size-Sector	None	0.88	0.96	0.83	0.86	0.77	0.71
CBP	State-Size-Sector	<10	0.88	0.97	0.84	0.97	0.82	0.96
CBP	State-Size-Sector	<10, >1000	0.92	0.97	0.91	0.97	0.87	0.96
CBP	County	None	0.99	0.98	1.00	0.99	1.00	0.99
CBP	Cty.-Size-Sector	None		0.94		0.87		0.73
CBP	Cty.-Size-Sector	<10		0.96		0.97		0.96
CBP	Cty.-Size-Sector	<10, >1000		0.96		0.97		0.96
CBP	Zip-Size-Sector	None		0.89		0.81		0.65
CBP	Zip-Size-Sector	<10		0.80		0.83		0.84
CBP	Zip-Size-Sector	<10, >1000		0.80		0.83		0.83

Source: NETS, CBP, and QCEW.

Notes: Simple correlations of cell-level employment and establishment counts, NETS and official sources. Exclusions refer to excluded establishment size classes. QCEW comparisons use NETS sample restricted to QCEW scope. CBP comparisons use NETS sample restricted to CBP scope.

Table 4: Cell-based correlations

industry measurement, yet size is not always easy to pin down given the employer versus nonemployer distinctions discussed above. Moreover, seasonal business fluctuations can easily result in the movement of establishments across size bins, and NETS measurement timing is heavily vulnerable to seasonality. In any case, we do not interpret the correlation gap between county aggregates and county-size-industry aggregates as necessarily indicating quality problems in NETS (aside from our seasonality concern); a significant portion of this discrepancy may be attributable to benign differences in labeling.

Table 4 suggests that NETS is in general agreement with official sources in terms of the distribution of economic activity across states, counties, establishment sizes, and industries when it is restricted to exclude the smallest establishments. Considerable mismeasurement of small establishments is a definite problem in NETS—indeed, a much bigger problem than the well-known challenge of measuring small establishments in official sources. But our broad conclusion is that NETS can be made reasonably representative of the U.S. econ-

Sample	Share Covered	
	Emp.	Estab.
Sample 1: ex. less than 10	0.848	0.223
Sample 2: ex. less than 10 & greater than 1000	0.743	0.223
Sample 3: ex. less than 10, greater than 1000 & NAICS 61	0.727	0.219
Memo: Private QCEW levels	113,326,720	8,994,650

Source: QCEW

Note: Share of QCEW employment accounted for by scope restrictions. 2014 data.

Table 5: QCEW coverage of industry scope restrictions

omy in terms of location, industry, and establishment size. This lends credibility to NETS-based studies of static business-level activity (in related work in progress, we find dynamic comparisons to be less appropriate). We therefore recommend that researchers ensure that NETS-based results are robust to sample restrictions that omit small establishments—and perhaps the largest establishments and the education sector as well. An important question is how costly these sample restrictions are in terms of coverage. Table 5 reports the fraction of QCEW establishments and employment that fall within our restricted samples.

Sample 1, which excludes small establishments and closes most of the gap between NETS and QCEW, still covers 85 percent of QCEW employment. Even the most restrictive sample, which excludes small and large establishments and the education sector, still covers 73 percent of workers. The coverage of establishments is, of course, much lower, since small establishments are much more common than large ones. Low establishment coverage may be a significant problem in some applications, but employment is typically the more important target of study given both its use as a measure of economic activity and the problems with counting establishments even in official data. The high employment share covered by our restricted samples is reassuring about the usefulness of NETS.<sup>24</sup>

<sup>24</sup>For comparison, another commonly used source of publicly available business microdata is Compustat, which reports detailed data on the universe of publicly traded businesses; the firms covered by Compustat account for well less than half of U.S. private sector employment (Davis et al. (2007)).

### 3.9 Data timing

As noted above, official sources record data in mid-March, while D&B data are collected continually throughout the year. The NETS sample for year  $t$  reflects a snapshot of the database in January of  $t$ . Therefore, for example, an employment number recorded by D&B in February of 2005 will appear in NETS as the 2006 observation. For this reason, some researchers have rolled NETS data back one year when comparing against official sources (that is, setting  $t^* = t - 1$ ).

We have not modified the NETS year data in our reported analysis for two reasons. First, the appropriate timing for comparison is actually not obvious. For example, if D&B collects an establishment's employment data in November 2005, then NETS will count these data as the 2006 observation. Our current approach would compare this November 2005 employment number in NETS to the March 2006 observations in official sources, for a timing difference of four months. Those researchers who roll NETS data back one year will compare the November 2005 NETS observation to the March 2005 observations in official sources, for a timing difference of eight months. It is true, however, that if D&B recording is uniformly distributed across the year, our method of leaving NETS year data unaltered will result in somewhat more error (on average) than a method that rolls NETS years back. But given potential lags in NETS recording and reporting, we do not see this difference as likely to be significant.

Second, in unreported results we performed all analyses in this paper with NETS data rolled back by one year (this means that instead of comparing NETS observations for 2000, 2007, and 2014 to the same years' data in official sources, we compare NETS observations that are reported as covering 2000, 2007, and 2014 to the years 1999, 2006, and 2013, respectively, in official sources). In a majority of comparisons, the relationship between NETS and official sources is closer when we leave NETS years unaltered than when we roll NETS years back. Comparisons by establishment size are almost uniformly better with unaltered NETS years. Comparisons by industry are more mixed, with a modest majority favoring unaltered

years, and in many cases the unaltered data are better by wide margins (while the margins for cases in which altered data are better tend to be small). Narrow-cell correlations in the unaltered-year data are higher in a large majority of cases. A few charts look visually better in the altered-year data, but most are visually similar.

We therefore prefer leaving NETS year data unaltered and interpreting NETS years at face value when comparing with official sources. However, other researchers may have needs for which the alternative approach is more appropriate.

## 4 Conclusion

We document a number of limitations with NETS. Coverage of very small establishments diverges markedly from official data sources, with significant apparent mismeasurement due to imputation. The small-establishment problem worsens significantly during the 2000s, and researchers must be careful to not interpret rising establishment counts in NETS over the last 20 years as indicative of robust business formation activity.<sup>25</sup> The NETS universe began the early 2000s being moderately too large to match the official employer universe, and by the early 2010s the NETS universe is far larger than the official employer universe but still smaller than the total employer and nonemployer universe by Census (and IRS) reckoning. In other words, differences between NETS and official sources do not simply reflect the notion that D&B has a more comprehensive and complete universe of business activity.

These differences between NETS and official sources are not reason to reject the use of NETS in research, however. We first show that the NETS universe can be restricted to approximate the scope of official sources in terms of industry and wage-and-salary employment. We then show that remaining discrepancies between NETS and official sources are largely driven by differences among small establishments, where imputation is prevalent.

---

<sup>25</sup>To the contrary, overwhelming evidence exists indicating that employer business formation has declined in recent decades; see, e.g., [Decker et al. \(2014\)](#).

Further, we find large differences in the education sector, where restricting the NETS sample to match official sources is difficult (note that this is not really a strike against NETS but rather a labeling difficulty. More involved efforts to determine whether individual establishments are public or private would likely close this gap). While we do show that NETS fails to adequately capture some key recent developments in mining, construction, and manufacturing, the extent to which these problems reflect industry labeling is unclear. Correlations of county-level aggregates between NETS and Census data are strikingly high. Moreover, at the narrow cell level (geography by industry by establishment size), we find correlations between NETS and official sources that are reassuringly strong in appropriately restricted samples. Our recommendation is that NETS users ensure results are robust to these sample restrictions. In related work in progress, we find stronger reason for caution when studying dynamic elements of NETS, consistent with previous state-level research. Our view is that static study is likely to be much safer than study of business dynamics. Note that we have not investigated the revenue data in NETS, which in principle could also be done by comparing NETS with public Census sources.

It is also worth noting that the high-quality microdata underpinning CBP, QCEW, and other official datasets are available through the continually expanding Federal Statistical Research Data Center (FSRDC) network and the BLS's visiting researcher program. These programs allow approved researchers to access data at least as detailed as what are available in NETS. Participating researchers must first submit proposals describing their application of the data, travel to a location where the data can be accessed securely, and subject all results to a disclosure process to ensure no sensitive information is released. Much productive research has been published through this process.<sup>26</sup> NETS has the advantage that the only barrier to access is the subscription fee; it can be used freely on any machine, and output does not have to undergo a review process; in this sense NETS is an appealing option.

We tend to favor the use of Census and BLS data when possible. Additionally, when ag-

---

<sup>26</sup>For examples, see the Center for Economic Studies working paper series at <https://ideas.repec.org/s/cen/wpaper.html>, and the many papers written using confidential BLS data.

gregated data are appropriate, official sources remain the gold standard.<sup>27</sup> However, when microdata are needed and speed or flexibility of analysis are requirements, NETS can be a tremendously useful resource.

## References

**Acs, Zoltan, William Parsons, and Spencer Tracy**, “High-Impact Firms: Gazelles Revisited,” Technical Report, Small Business Administration, Office of Advocacy 2008.

**Amezcuca, Alejandro S.**, “Boon or Boondoggle? Business Incubation as Entrepreneurship Policy.” PhD dissertation, Syracuse University 2010.

**Bayard, Kimberly, Ryan A. Decker, and Charles Gilbert**, “Natural Disasters and the Measurement of Industrial Production: Hurricane Harvey, a Case Study,” *FEDS Notes*, 2017.

**Becker, Randy, Joel Elvery, Lucia Foster, C. J. Krizan, Sang Nguyen, and David Talanl**, “A Comparison of the Business Registers Used by the Bureau of Labor Statistics and the Bureau of the Census,” Technical Report 2005.

**Choi, Taelim, Anil Rupasingha, John C Robertson, and Nancey Leigh**, “The Effects of High Growth on New Business Survival,” *The Review of Regional Studies*, 01 2017, 47, 1–23.

—, **John C. Robertson, and Anil Rupasingha**, “High-growth firms in Georgia,” FRB Atlanta Working Paper 2013-20, Federal Reserve Bank of Atlanta December 2013.

---

<sup>27</sup>As noted previously, collectors and producers of official data face considerable challenges as well, and a large literature (much of it written by expert employees of statistical agencies) documents shortcomings of official sources. Yet these sources are simply unrivaled in terms of conceptual consistency and scientifically driven measurement due to the explicit measurement goals of their data collection processes and to the professional and trained staff undertaking the data collection and production; in addition, the statistical agencies are continually engaged in interactions with the business and academic communities to improve the measurement process. Moreover, as we show in the appendix, the U.S. is fortunate to have two independently constructed business registers that, while not in complete harmony, are remarkably consistent and indeed more consistent than is NETS with either, supporting the case for their accuracy. In this respect, we view NETS as a complement (rather than a substitute) to official sources.

- Crane, Leland D., Ryan A. Decker, and Bo Yeon Jang**, “Flood Effects,” Unpublished Draft, Board of Governors of the Federal Reserve System 2017.
- Cromwell, Erich W.**, “Topics on Labor and Public Policy.” PhD dissertation, Florida State University 2015.
- Currie, Janet, Stefano DellaVigna, Enrico Moretti, and Vikram Pathania**, “The Effect of Fast Food Restaurants on Obesity and Weight Gain,” *American Economic Journal: Economic Policy*, August 2010, 2 (3), 32–63.
- Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda**, “Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms,” in “NBER Macroeconomics Annual 2006, Volume 21” NBER Chapters, National Bureau of Economic Research, Inc, 2007, pp. 107–180.
- , —, —, **Ron S. Jarmin, C. J. Krizan, Javier Miranda, Alfred Nucci, and Kristin Sandusky**, “Measuring the Dynamics of Young and Small Businesses: Integrating the Employer and Nonemployer Universes,” in “Producer Dynamics: New Evidence from Micro Data,” National Bureau of Economic Research, 2009, pp. 329–366.
- Decker, Ryan, John Haltiwanger, Ron S. Jarmin, and Javier Miranda**, “The Role of Entrepreneurship in US Job Creation and Economic Dynamism,” *Journal of Economic Perspectives*, 2014, 28 (3), 3–24.
- DeSalvo, Bethany, Frank Limehouse, and Shawn D. Klimek**, “Documenting the Business Register and Related Economic Business Data,” Technical Report, Center for Economic Studies working paper CES-WP-16-17 2016.
- Donegan, Mary**, “Inside the Triangle: Does Database Selection Alter our Understanding of Urban Industrial Systems?,” Technical Report 2014.
- Echeverri-Carroll, Elsie and Maryann Feldman**, “Chasing Entrepreneurial Firms,” Technical Report 2017.



**Fairman, Kristin, Lucia Foster, C.J.Krizan, and Ian Rucker**, “An Analysis of Key Differences in Micro Data: Results from the Business List Comparison Project,” Technical Report, Center for Economic Studies working paper CES 08-28 2008.

**Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen**, “Do Credit Market Shocks affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and ‘Normal’ Economic Times,” Working Paper 20704, National Bureau of Economic Research November 2014.

— and —, “Do Credit Market Shocks affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and ‘Normal’ Economic Times,” 2012.

**Groizard, Jose L., Priya Ranjan, and Antonio Rodriguez-Lopez**, “Trade Costs and Job Flows: Evidence from Establishment-Level Data,” *Economic Inquiry*, 2015.

**Haltiwanger, John, Ron S. Jarmin, and Javier Miranda**, “Who Creates Jobs? Small versus Large versus Young,” *The Review of Economics and Statistics*, 2013, 95 (2), 347–361.

**Isenberg, Emily, Liana Christin Landivar, and Esther Mezey**, “A Comparison Of Person-Reported Industry To Employer-Reported Industry In Survey And Administrative Data,” Working Papers 13-47, Center for Economic Studies, U.S. Census Bureau September 2013.

**Jarmin, Ron S. and Javier Miranda**, “The Longitudinal Business Database,” Technical Report, Center for Economic Studies 2002.

**Kolko, Jed and David Neumark**, “Do Enterprise Zones Create Jobs? Evidence from California’s Enterprise Zone Program,” *Journal of Urban Economics*, 2010, 68 (1).

**Mach, Traci L. and John D. Wolken**, “Examining the Impact of Credit Access on Small Firm Survivability,” in “Small Businesses in the Aftermath of the Crisis: International Analyses and Policies,” Heidelberg: Physica-Verlag HD, 2012, pp. 189–210.

**Neumark, David, Brandon Wall, and Junfu Zhang**, “Do Small Businesses Create More Jobs? New Evidence for the United States from the National Establishment Time Series,” *The Review of Economics and Statistics*, 2011, 93 (1), 16–29.

—, **Junfu Zhang, and Brandon Wall**, “Employment Dynamics and Business Relocation: New Evidence from the National Establishment Time Series,” Working Paper 11647, National Bureau of Economic Research October 2005.

**Walls & Associates**, “National Establishment Time-Series (NETS) Database,” 2014.

**Walls, Donald**, “Understanding Data in the NETS Database,” *NETS Documentation*, 2008.

## A Appendix: CBP and QCEW coverage

### A.1 CBP coverage

CBP covers private business establishments, with the following exclusions:<sup>28</sup>

- NAICS 111 and 112 (Crop and animal production)
- NAICS 482 (Rail transportation)
- NAICS 491 (Postal service)
- NAICS 525110, 525120 and 525190 (Health, welfare and vacation funds)
- NAICS 525920 (Trusts, estates and agency accounts)
- NAICS 814 (Private households)
- NAICS 92 (Public administration)

CBP also *includes* government-run establishments in the following industries:

- NAICS 4248 (Government sponsored wholesale liquor establishments)
- NAICS 44531 (Retail liquor stores)
- NAICS 511130 (Book publishers)
- NAICS 522120 (Federally-chartered savings institutions)
- NAICS 522130 (Federally-chartered credit unions)
- NAICS 662 (Hospitals)

---

<sup>28</sup>The material in this section is from <https://www.census.gov/programs-surveys/cbp/technical-documentation/methodology.html>

## A.2 QCEW coverage

We use QCEW tables that exclude federal, state, and local government establishments. The following industries are partially or entirely excluded from QCEW:<sup>29</sup>

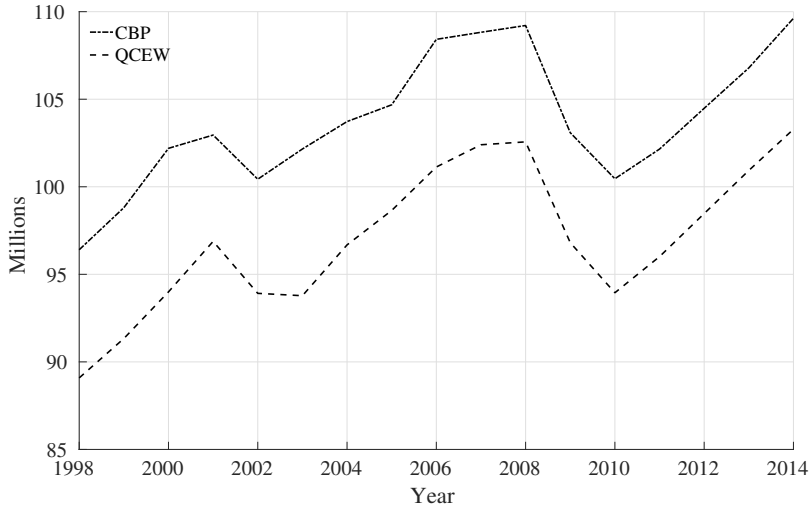
- NAICS 11 (Agriculture etc.)
- NAICS 482 (Railroads)
- NAICS 813 (Religious groups)
- NAICS 814 (Domestic workers)

When we compare QCEW to CBP, we exclude all 4-digit NAICS which are completely or partially excluded from either dataset.

---

<sup>29</sup>Based on Table A of <https://www.bls.gov/cew/cewbultn15.htm#Employment>

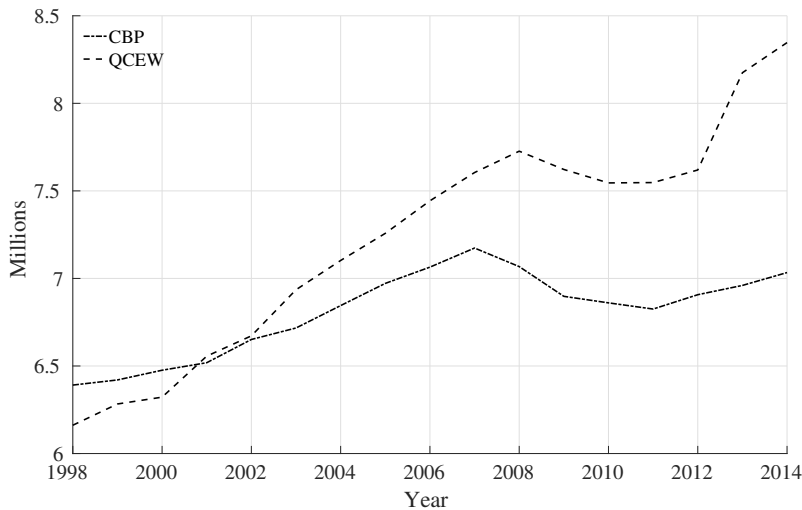
## B Additional Tables and Figures



Source: CBP, QCEW.

Note: Starting from the CBP and the private-sector QCEW, we exclude the following NAICS, which are (partially) out of scope for either CBP or QCEW: 11, 814, 4248, 4453, 5111, 5221, 622, 4821, 8131

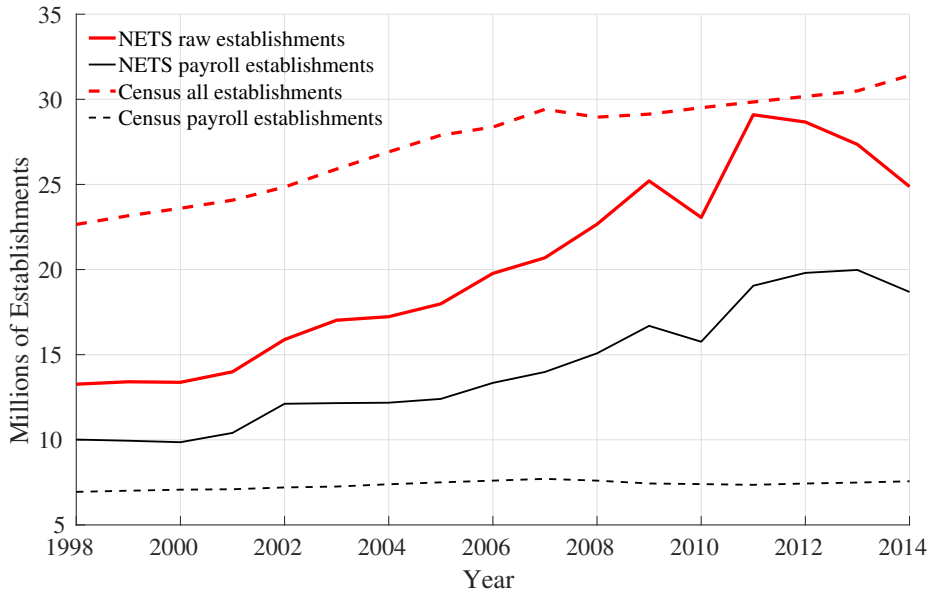
Figure 5: Aggregate employment, CBP versus QCEW (intersection of scopes)



Source: CBP, QCEW.

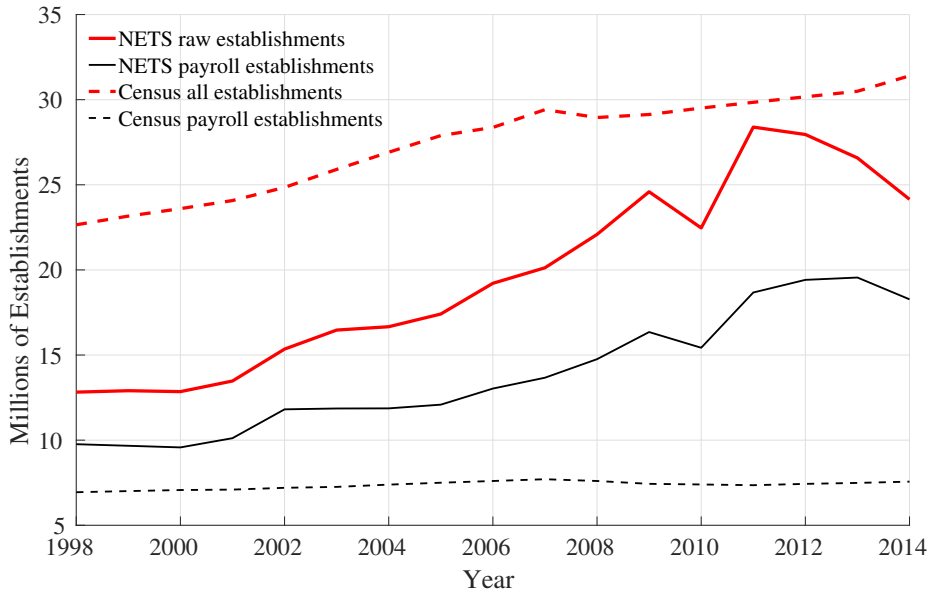
Note: Starting from the CBP and the private-sector QCEW, we exclude the following NAICS, which are (partially) out of scope for either CBP or QCEW: 11, 814, 4248, 4453, 5111, 5221, 622, 4821, 8131

Figure 6: Aggregate establishment counts, CBP versus QCEW (intersection of scopes)



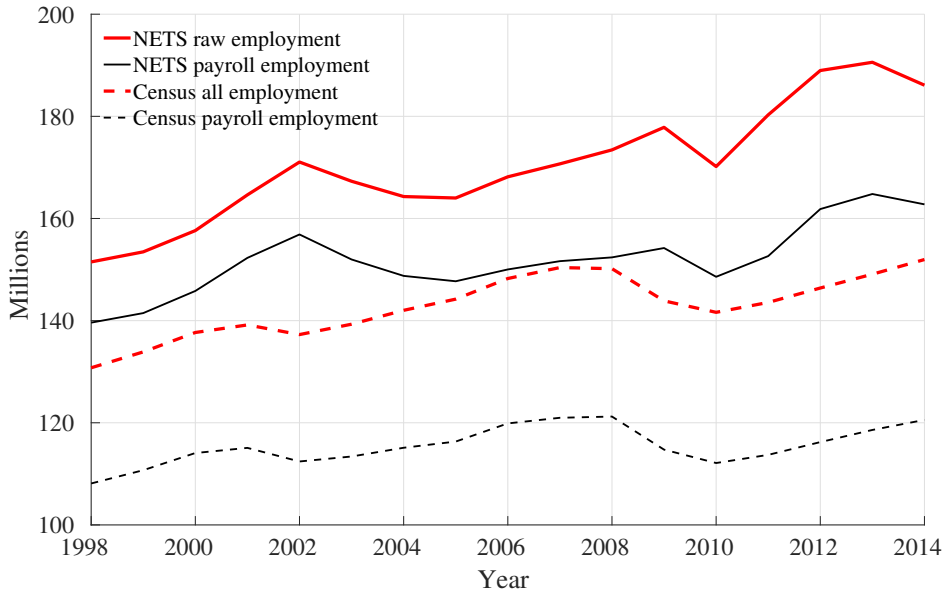
Source: NETS database, County Business Patterns, Census Nonemployer Statistics.  
 Note: NETS sample not restricted to CBP scope.

Figure 7: Aggregate establishment counts (including NETS government establishments)



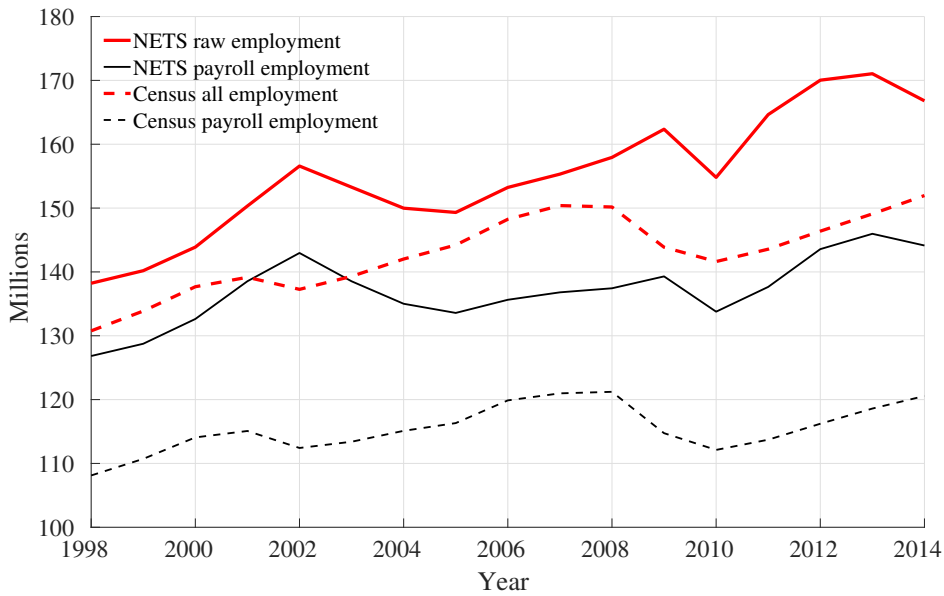
Source: NETS database, County Business Patterns, Census Nonemployer Statistics.  
 Note: NETS sample restricted to CBP scope.

Figure 8: Aggregate establishment counts (separate NETS lines of business)



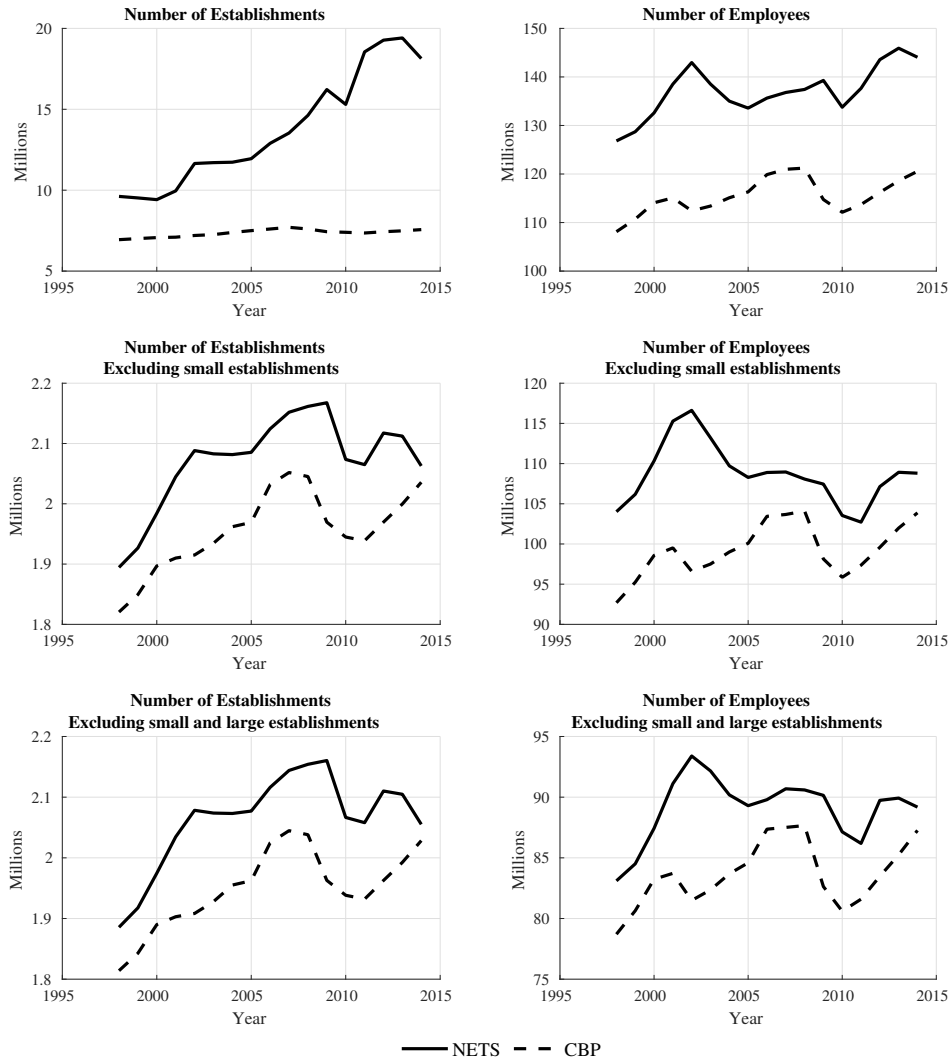
Source: NETS database, County Business Patterns, Census Nonemployer Statistics.  
 Note: NETS sample not restricted to CBP scope.

Figure 9: Aggregate employment (including NETS government establishments)



Source: NETS database, County Business Patterns, Census Nonemployer Statistics.  
 Note: NETS sample restricted to CBP scope.

Figure 10: Aggregate employment (separate NETS lines of business)

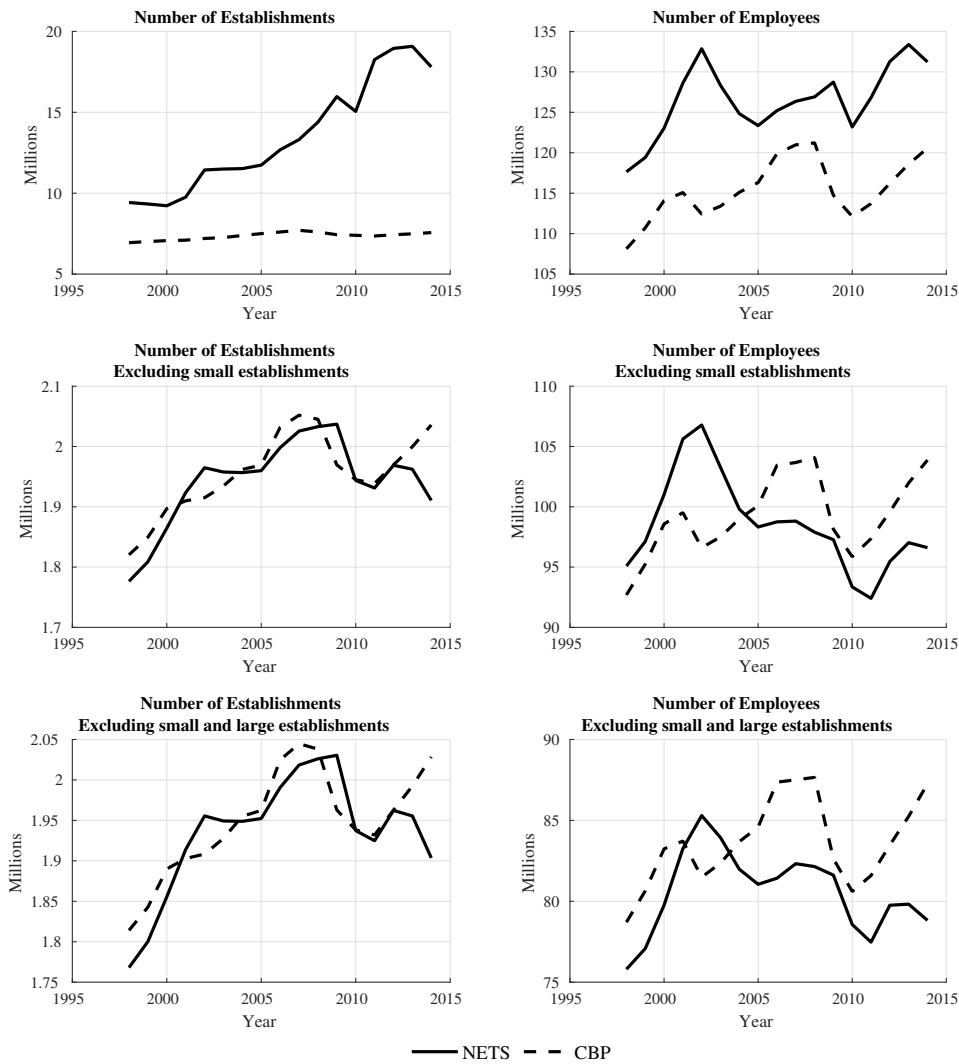


Source: NETS, CBP

Notes: NETS sample is restricted to CBP scope (NET sample baseline-CBP). "Small" establishments are those with less than 10 employees. "Large" establishments are those with 1000 or more employees.

Figure 11: NETS versus CBP





Source: NETS, CBP

Notes: NETS sample is restricted to CBP scope, and NAICS 61 is excluded from both datasets (NET sample ex. 61-CBP). "Small" establishments are those with less than 10 employees. "Large" establishments are those with 1000 or more employees.

Figure 12: NETS versus CBP excluding educational services

Size class	2000			2007			2014		
	QCEW	CBP	Pct. diff.	QCEW	CBP	Pct. diff.	QCEW	CBP	Pct. diff.
1 to 4	3439.5	3471.0	-0.91	4438.3	3929.7	12.94	5162.8	3849.4	34.12
5 to 9	1185.4	1242.9	-4.63	1305.5	1327.9	-1.69	1283.1	1273.4	0.76
10 to 19	791.9	825.1	-4.03	872.8	905.5	-3.61	901.4	895.6	0.65
20 to 49	559.1	574.6	-2.69	616.7	626.7	-1.59	626.3	632.0	-0.90
50 to 99	195.8	200.3	-2.25	210.7	215.8	-2.39	212.9	214.8	-0.88
100 to 249	111.2	117.0	-4.96	118.7	122.3	-2.94	118.4	122.0	-3.02
250 to 499	27.1	29.8	-9.13	28.4	30.1	-5.51	28.6	30.3	-5.86
500 to 999	8.4	10.2	-17.85	9.7	10.5	-7.38	9.3	10.5	-11.01
1000+	3.8	5.0	-24.27	3.9	5.1	-23.18	4.0	5.4	-26.35

Source: CBP, QCEW

Notes: Levels are in thousands. Percent differences are CBP less QCEW divided by CBP establishment counts. Both CBP and QCEW are restricted to the intersection of their scopes. Starting from the CBP and the private-sector QCEW, we exclude the following NAICS, which are (partially) out of scope for either CBP or QCEW: 11, 814, 4248, 4453, 5111, 5221, 622, 4821, 8131.

Table 6: CBP versus QCEW, establishment counts by size class

Size class	2000			2007			2014		
	QCEW	CBP	Pct. diff.	QCEW	CBP	Pct. diff.	QCEW	CBP	Pct. diff.
1 to 4	5.5	5.9	-7.07	6.8	7.0	-3.12	7.4	6.9	6.85
5 to 9	7.8	8.2	-4.58	8.6	8.8	-1.84	8.5	8.5	0.67
10 to 19	10.7	11.1	-3.93	11.8	12.2	-3.38	12.2	12.1	0.82
20 to 49	16.9	17.4	-2.57	18.6	18.9	-1.39	18.9	19.1	-0.97
50 to 99	13.5	13.8	-2.16	14.5	14.9	-2.57	14.6	14.7	-0.85
100 to 249	16.7	17.6	-5.35	17.8	18.2	-2.38	17.6	18.2	-3.00
250 to 499	9.2	10.1	-9.23	9.7	10.2	-5.14	9.7	10.4	-6.67
500 to 999	5.7	6.9	-17.69	6.6	7.1	-7.48	6.3	7.1	-11.54
1000+	8.0	11.1	-28.13	8.0	11.5	-30.35	8.0	12.6	-36.34

Source: CBP, QCEW

Notes: Levels are in millions. Percent differences are CBP less QCEW divided by CBP employment levels. Both CBP and QCEW are restricted to the intersection of their scopes. Starting from the CBP and the private-sector QCEW, we exclude the following NAICS, which are (partially) out of scope for either CBP or QCEW: 11, 814, 4248, 4453, 5111, 5221, 622, 4821, 8131.

Table 7: CBP versus QCEW, employment by size class

Size class (Employees)	Percent difference					
	2000		2007		2014	
	Emp.	Estab.	Emp.	Estab.	Emp.	Estab.
1 to 4	99.10	59.02	123.77	105.30	174.77	156.24
5 to 9	23.67	23.07	22.90	21.62	45.11	42.17
10 to 19	10.14	8.52	9.48	8.20	0.27	-0.29
20 to 49	11.25	9.18	9.92	7.46	2.15	-0.26
50 to 99	12.13	11.81	10.96	10.55	12.62	10.76
100 to 249	4.42	5.73	1.15	3.10	6.81	8.10
250 to 499	9.02	9.72	4.31	5.54	3.74	4.01
500 to 999	32.51	33.14	7.62	8.88	6.77	7.42
1000+	100.59	78.82	56.40	39.39	64.18	35.30
Aggregate	26.85	38.74	20.99	66.61	26.33	102.76
Aggregate ex. less than 10 employees	22.48	9.29	13.20	7.95	12.69	1.68
Aggregate ex. less than 10 and greater than 1000 employees	11.25	9.09	7.18	7.86	5.38	1.58

Source: NETS, QCEW.

Note: Difference between NETS and QCEW employment as percent of QCEW employment. NETS sample restricted to QCEW scope.

Table 8: NETS versus QCEW by establishment size

Industry	Percent difference					
	2000		2007		2014	
	Excl <10	All	Excl <10	All	Excl <10	All
11 Ag., For., Fish., Hunt	-99	-99	-98	-99	-99	-99
21 Mining	54	54	6	10	-14	-12
22 Utilities	-34	-32	-30	-26	-36	-31
23 Construction	-3	6	-5	4	7	23
31-33 Manufacturing	25	27	34	37	43	47
42 Wholesale Trade	19	24	13	18	6	13
44-45 Retail Trade	-13	-3	-9	2	-1	7
48-49 Trans., Warehous.	2	6	-6	1	-18	-6
51 Information	15	20	30	37	39	47
52 Finance, Insurance	35	34	15	20	18	23
53 Real Est., Rent., Leas.	65	67	76	82	74	96
54 Prof., Sci., Tech. Svcs	39	41	24	29	14	25
55 Management	-85	-84	-84	-81	-79	-74
56 Admin., Waste Mgmt	-21	-14	-24	-0	-29	20
61 Education Svcs	441	422	364	347	379	371
62 Health, Social Asst.	15	18	2	8	-8	0
71 Arts, Entertain., Rec.	38	50	29	44	16	38
72 Accom., Food Svcs	-1	2	-5	3	-10	3
81 Other Svcs	-30	-19	-34	-23	-42	-26

Source: NETS and QCEW

Notes: Difference between NETS and QCEW employment as percent of QCEW employment by NAICS sector. NETS sample restricted to QCEW scope.

Table 9: NETS versus QCEW by sector