Estimating the U.S. Transgender Population: Evidence from Federal Administrative Records

Benjamin Cerf Harris*
U.S. Census Bureau

August 2013

DRAFT. DO NOT CITE OR QUOTE WITHOUT PERMISSION.

Abstract

This paper uses Social Security Administration (SSA) records of individuals' first names and sex-coding to estimate the U.S. adult transgender population. Estimates are based on observed changes in how individuals' first names and sex-coding change in the SSA administrative records over time. I present two estimates: the more conservative estimate derives from coincident changes in sex-coding from male to female (or female to male) and first name from traditionally male to traditionally female (or vice versa). Because the SSA requires individuals to undergo genital sexual reassignment surgery (SRS) to change the sex-coding on their records, this estimate only identifies transgender people who have undergone such surgery. Although it yields estimates that are comparable to most studies on the prevalence of transsexualism in other countries, this measure is somewhat problematic since not all transgender people choose or can afford SRS. Therefore, a more inclusive estimate counts individuals who have changed their first name from high-probability female to high-probability male (or vice versa) but also includes transgender people who have not undergone SRS. My estimates suggest that as of 1 April 2010 17,481 to 24,784 adult transgender people in the U.S. underwent SRS. However 126,679 to 359,274 adults legally changed their names in ways that are consistent with gender transition, suggesting that those who transition via SRS make up only a small fraction of the transgender population. Furthermore, I find substantial heterogeneity in the paths by which people mark their transition legally. The majority of people who change both their name and their sex-coding do so simultaneously, although approximately 20% change their name first and their sex-coding about 5–6 years later on average. Finally, I explore the historic trends in transgender-consistent claims approved by the SSA. I find evidence of transgender-consistent claims as early as the 1930s and historic changes in the rates of transgender-consistent claims that correspond to major events in transgender American history. It is important to note that there are several types of transgender people my estimation strategy does not identify (e.g., people without SSNs), and so I am likely *underestimating* the transgender population. Nevertheless, this research demonstrates the potential of administrative records to address important questions that conventional survey data cannot answer.

Keywords: Administrative Records, Gender Dysphoria, Gender Identity, Record Linkage, Transgender, Transsexual

---

Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

1 Introduction

In recent years, transgender issues and images have been increasingly present in popular media, literature, journalism, and the legal landscape of the United States. The popular television show, *Glee,* featured a transgender high school student in 2012; the 2013 series, *Orange is the New Black,* alerted viewers to unique issues of race and gender identity that inmates and their families face by featuring a black transgender prisoner played by transgender actor Laverne Cox; and in the same year, retired Navy SEAL, Kristin Beck (formerly Chris), published her memoir about coming out as a transgender woman after 20 years of service (Votta, 2012; Ryan, 2013; Stanglin, 2013). At the same time, California passed a 2013 law protecting the rights of transgender elementary and high school students to utilize restrooms and participate in sports teams that reflect their asserted gender identities; Cook County Jail—and jails around the country—have begun to update their policies around transgender detainees; and a $1.35 million grant was recently opened to assess how to allow transgender soldiers to serve openly in the military (Lovett, 2013; Lu, 2011; Sosin, 2013). Despite increasing awareness of transgender issues and numerous qualitative studies on transgender experiences, there is very little quantitative evidence about the social and economic lives of transgender people in the United States. In fact, we have been unable to answer satisfactorily even the most fundamental question, "How many transgender people reside in the United States?" The main challenge—but certainly not the only challenge—is one of data. Random sample surveys are not feasible, since sample sizes would need to be impractically large to ensure adequate coverage of the transgender population.[1] Something closer to a census is needed, although questions about gender identity are not included in the nation's decennial census. Another important challenge, however, is definitional. Even if there were a large enough survey to ensure a large random sample of transgender people, it is not clear what question (or questions) would address the various understandings of the word 'transgender' or the many incarnations that gender identity takes once we abandon the framework of binary sex categories.

This paper offers a unique approach to estimating the transgender population in the U.S. using a large, confidential, national data source. Specifically, I analyze the Social Security Administration (SSA) Numeric Identification System (colloquially, the "Numident") to identify all adult Social Security Number (SSN) holders who permanently change their first names from traditionally male to traditionally female names (or vice versa) and also permanently change the sex-coding associated with their SSN in a the same direction. This methodology allows me to estimate the transgender population according to some (but not nearly all) of the various definitions of what it means to be transgender. Until very recently the SSA required evidence of genital sexual reassignment surgery (SRS) for a person to change the sex-coding on his or her records, so looking at coincident changes in names and sex-coding corresponds to a definition of transgenderism that depends on surgical intervention.[2] However, looking at name changes only corresponds to a more inclusive definition that depends more on how a person presents gender in society, regardless of the particular path that person took to realize that presentation of gender. Using these two main estimates, I explore the following three research questions: First, how many adult SSN holders who were alive on 1 April 2010 (census day) changed their first names, their sex-coding, or both in ways that are consistent with gender transition? Second, what are the main event-paths associated with these changes? Do most people only change their names, or do most change both their names and their sex-coding? Of those who change both their name and their sex-coding, do they typically change their name first and their sex-coding later, or do they typically change both their name and their sex-coding simultaneously? Finally, to what extent have the rates of transgender-consistent name and sex-coding changes grown or decreased since the creation of the

[1] A notable exception is the Massachusetts Behavioral Risk Factor Surveillance System Survey, which is a telephone survey that began to include a question about transgender identity in 2007. Conron *et al.* (2012) pool several years of the survey to obtain estimates of the transgender population in the state; 0.5 percent of their sample identified as transgender.

[2] Transgender people who undergo SRS are sometimes referred to as 'transsexuals', although not all transsexual people undergo genital SRS.

SSA in 1936?

I find that the ways we define 'transgender' has a large bearing on the answer to the question about the size of the transgender population. Depending on the strength of the rules used for assigning genders to names, I find that in 2010 there were 17,481–24,784 adult SSN holders who changed both their first names and their sex-coding in ways that are consistent with gender transition. However, a more inclusive definition of 'transgender', which depends solely on gender being asserted through a person's legal first name, paints a very different picture. Applying this definition shows that 126,679–359,274 adult SSN holders alive in 2010 had changed their first names from traditionally male to traditionally female names or vice versa. These numbers suggest that 0.1 percent of the adult population is transgender.Of those who change their names and their sex-coding, 12,320–16,755 appear to have transitioned from male to female (MTF), and 5,161–8,029 appear to have transitioned from female to male (FTM).[3]

The large disparity between estimates helps address the second research question. In particular, it shows that, for whatever reason, most transgender people do not undergo SRS. Approximately 76.2–93.1 percent change their name but not their sex-coding. Of those 6.9–13.8 percent who do change both their names as well as their sex-coding, about 80 percent make both changes to their SSA records simultaneously. Nearly all others change their name first and their sex-coding later. For both MTF and FTM transgender people, these changes occur primarily when people are in their mid- to late-thirties; and of those who change their names before their sex-coding, the average lag between the name change and the sex-coding change is 5.2–6.9 years.

To investigate the third research question, I compare annual rates of transgender-consistent name and sex coding changes, or *claims,* to the rates of all other types of claims (e.g., surname changes, corrections to dates of birth, etc.) since the SSA was created in 1936. I identify several noteworthy trends in the data. For example, there is strong evidence of transgender-consistent name and sex-coding changes even during the earliest years of the SSA. These claims, as well as all other types of claims, steadily increased in frequency until the 1970s, when the number of claims per year increased dramatically. For everyone, the 1970s was a period characterized by increasing legislation requiring the use of the SSN as a unique identifier, and so it is no wonder that more claims would start to be filed. However, the 1970s was also the decade that featured some early public and legal advances for transgender rights in the United States, including a Connecticut district court ruling that ensured a transgender person's right to change the information on her birth certificate as well as the implementation of the first city ordinance protecting the civil rights of transgender people in Minneapolis. Perhaps not surprisingly, then, growth in transgender-consistent claims in the 1970s outpaced the already rapid growth in all other types of claims. Later, in 2002, at the advent of the War on Terror, increased concern about verifying identity led the Department of Homeland Security (DHS), Internal Revenue Service (IRS), and the SSA to collaborate in checking employee information supplied to the IRS by employers against the information in the SSA Numident. If transgender employees had not legally changed their name or sex-coding with the SSA, then their employers would receive a letter indicating as much and effectively disclosing the employee's transgender status. These letters increased the incentive of people who had undergone SRS but had not updated their information with the SSA to do so, and indeed during the early 2000s, transgender-consistent name changes decreased while transgender-consistent sex-coding changes increased.

There are several limitations to this research design, and readers should interpret the results with the following caveats in mind. Although I am using population-based data, these are estimates, and are subject to many types of measurement error other than sampling error. First, because my data only reflect the population of SSN holders, I am unable to identify transgender people who immigrated to the United States and do not

[3]These figures imply a prevalence of 8.2–11.15 MTF transgender people out of 100,000 and a prevalence of 3.26–5.07 FTM transgender people out of 100,000. These estimates are quite similar to estimates attained by surveying doctors and psychologists in other countries about the number of patients who seek SRS (See Bakker *et al.*, 1993; De Cuypere *et al.*, 2007; Wilson *et al.*, 1999, for studies in the Netherlands, Belgium, and Scotland, respectively).

have an SSN and people who transitioned before acquiring an SSN. Beyond that, transgender people whose original or new names are not as strongly male or female will not be counted, nor will transgender people who do not change their names. People may not change their names because their original name is preferred (for example, if the original name is gender-neutral), or because the process of legally changing a name is time-consuming, expensive, or made difficult due to administrative barriers such as those outlined in Spade (2008). It is important to note that the financial and time costs associated with legal name changes also vary substantially across states. Inconsistent application of SSA policies by individual clerks may also prohibit some transgender people from registering their name or sex-maker change; and people may change their names through common-law, rather than through a formal process, and these people would not be identified (Spade, 2008). Furthermore, to the extent that poverty and financial hardship is correlated with race and ethnicity, my estimates may underrepresent transgender people who are also members of minority or otherwise disadvantaged and marginalized groups. Second, in order to make my results comparable with previous work, I only consider name and sex-coding changes that occur after a person turns 16 years old, however it is possible and increasingly common for children to transition prior to their 16th birthday with parental consent (Dvorak, 2012; Spiegel, 2008). Therefore, my estimation strategy misses anyone who transitions prior to turning 16. By design, it also would not identify transgender or gender dysphoric people who have not taken any steps to transition. In addition, my methodology does not address appropriately the situations of some *intersex* people. People whose bodies were not easily categorized as male or female may have undergone surgical assignment of a particular sex shortly after birth; if those people later grew up to reject the decisions made by parents and physicians, then they may be wrongly identified as transgender by my estimation strategy (Meier and Labuski, 2013). Finally, people whose names or sex-coding that were changed erroneously, but not yet corrected, will by wrongly categorized as transgender. So long as this group does not outweigh those who are transgender but not counted by my estimation strategy, my estimates should be interpreted as *under*estimates.

With these caveats in mind, this paper makes several key contributions to the literature investigating the size of the transgender population. My first contributions are methodological. By far the most common research design to this point has been to count the number of sexual reassignment surgeries performed in countries with national medical registers (Bakker *et al.*, 1993; Landén *et al.*, 1996; Reed *et al.*, 2009) or to survey (or take a census of) surgeons about how many such surgeries they had performed (De Cuypere *et al.*, 2007; Pauly, 1968; Tsoi, 1988; Wålinder, 1968; Wilson *et al.*, 1999). Such studies limit the number of people under consideration by identifying only those who underwent SRS, but also by identifying only those people whose economic circumstances are sufficient pay for the costly surgeries. My methodology has the desirable feature of producing comparable estimates to the broader literature because of the policies requiring SSN holders to provide evidence of having undergone SRS in order to change the sex-coding associated with their SSN; however the real strength of this methodology is that, by also estimating the number of people who change their names in ways that are consistent with gender transition but do not necessarily change their sex-coding, I am able to show how the subpopulation of people who receive SRS maps into the much larger population of transgender adults who do not receive SRS. To the best of my knowledge, this is the first study to use changes in first names to estimate the size of the transgender population.

This work also contributes to the growing literature demonstrating the benefits of using certain state, federal, and commercial administrative records to answer statistical questions that cannot be answered using existing survey and census data. Even as administrative records gain fanfare over their ability to augment, quality control, and even replace the products of statistical agencies (for example, in 2011 Statistics Austria initiated an entirely administrative record-based census), not nearly enough opportunities are seized to utilize administrative records to explore demographic, economic, and social phenomena that would be difficult or impossible to answer using conventional statistical data. Using administrative records (other than medical registers) to estimate the size of the transgender population is one such example. I am aware of only two other studies that use administrative records to estimate the transgender population. The first involves counting

the number of passport holders in New Zealand whose sex-coding was either removed entirely or changed, and the second involves counting the number court cases in which people applied for legal changes of name and sex status in Germany (Veale, 2008; Weitze and Osburg, 1996, respectively).

My other contributions are empirical. First, only two studies estimate the transgender population in the United States. Pauly (1968) used the number of applications for SRS submitted to the Johns Hopkins Gender Identity Clinic to estimate that 2,500 transsexual adults resided in the U.S. during the year of his study. Recent work by Gates (2011) places the estimate much higher. Gates combines multiple survey results from California and Massachusetts to estimate that 0.3 percent of the total United States population is transgender, which works out to 697,529 people.[4] This paper provides a much needed update to Pauly's (1968) estimates of the transsexual population and simultaneously provides more broadly defined estimates of the transgender population that can be compared with Gates's (2011) estimates. Because I analyze data that include *every* adult SSN holder who was alive on 1 April 2010, my estimates do not rely on the strong assumption that the proportion of the population that is transgender is the same in all states as it is in California and Massachusetts.

Finally, this is the first study to estimate not only the *stock* of transgender people at a given point in time, but also the *flows* of gender transitions continuously over nearly eighty years. These results will help illuminate the history of transgender experiences in the United States during the 20th and early 21st centuries, and they can help pave the way toward evaluating how policy changes influence the rate of transgender-consistent claims. For example, in June 2013, the SSA lifted its requirement that people seeking sex-coding changes must provide evidence of having undergone genital sexual reassignment surgery SSA (2013). The analysis presented here can act as a benchmark for evaluating the effect this recent policy change will have on the rate of sex-coding change claims submitted to the SSA in the future.

The next section gives background on how gender is documented in the SSA, discusses the data, and describes the estimation strategy in detail. Section 3 presents results for the several estimates of the U.S. transgender population, results on the various paths people take in documenting their gender transitions, as well as my analysis of how the rates of transgender-consistent claims within the SSA have changed over time. Finally, Section 4 reviews the limitations of the study, summarizes the main conclusions, and discusses directions for future research.

2 Analytic Framework

2.1 Background

It is very important for a person's documented identity to match the basic characteristics of that person's social and physical identity. Barriers to updating the sex-coding and name on a driver's license, birth certificate, passport, or social security card can lead to involuntary disclosure of transgender status to police, sales clerks, employers, and any of the many other people who ask for identification throughout daily life. This can in turn expose transgender people to discriminatory treatment and violence (Grant *et al.*, 2011; Spade, 2008). Thus, the incentives for transgender people to update their identity documentation are strong; however several state and federal institutions can make changing gender documentation extremely difficult and can lead to a single person's gender being inconsistently documented. In a comprehensive inventory of gender documentation policies across various record-keeping agencies at both the state and federal levels, Spade (2008) demonstrates this thoroughly. For example, a transgender man who has undergone hormone therapy

[4]Specifically, Gates combines the 2009 California Health Interview Survey estimates that 3.2 percent of Californians identify as lesbian, gay, or bisexual with the 2003 California LGBT Tobacco Survey estimates that 3.2 percent of lesbian, gay, bisexual, or transgender people identify as transgender, to estimate that 0.1 percent of the total California population is transgender. Gates then averages this with Conron *et al.*'s (2012) estimate from Massachusetts of 0.5 percent to get the final estimate of 0.3 percent.

and chest surgery, but not genital surgery, can change the sex-coding on his passport but not in his SSA records.

To change information associated with an SSN, a person must complete and submit a form SS-5 to the SSA. The SS-5 is used for surname changes, corrections to information on date of birth, corrections to erroneously entered sex-coding, and of course changes to first names and changes to sex-coding following gender transitions. Prior to 2002, individuals wishing to change their sex-coding to reflect a gender transition were required to submit with their SS-5 evidence of the transition in the form of a physician's letter indicating the person had undergone or was scheduled to undergo "complete" sexual reassignment surgery, which is typically interpreted as genital surgery. In late 2002, the SSA adopted the stricter requirement that the surgery needed to be completed prior to the application (Spade, 2008). This was the policy in place on 1 April 2010, although the SSA recently updated the policy and no longer requires surgery (SSA, 2013). I was unable to attain documentation on the policies governing the requirements for sex-coding changes for transgender people prior to the rule in place in early 2002, although it is likely that rules have become stricter over time. Puckett (2009) points out that at the beginning of the SSA, the information on a person's record was based exclusively on what that person asserted to be true.[5]

2.2 Data

The data used in this paper examines the administrative records of the Social Security Administration, provided to the U.S. Census Bureau under Titles 5, 13, and 42 of the U.S. Code.[6] The SSA Numerical Identification System (or Numident) is an administrative database containing the name, date of birth, and a sex-coding for every Social Security Number (SSN) holder. Furthermore, the Numident contains a record for every transaction, or claim, that changes the information associated with a given SSN. The U.S. Census Bureau acquires the Numident through a data sharing agreement with SSA, and uses the data to facilitate record linkage and statistical operations. To ensure non-disclosure of SSNs, the Census Bureau strips each SSN from its version of the Numident and replaces it with a Protected Identification Key (PIK), which is a unique indentifier used internally by the Census to link individuals across data sets for statistical and research purposes. In this study, information on date of death is attained by the publicly available SSA Death Master File (DMF) and merged on by PIK. In addition, the SSA sends quarterly updates to inform the Census of the creation of new SSNs as well as changes and corrections of information associated with existing SSNs. Thus, the Numident contains a record for the *population* of all SSN holders as of the most recent update. This study evaluates all the Numident records from 1936 to the end of 2010. I restrict my universe of analysis to those people who were alive and over 16 years old on 1 April 2010.[7]

Table 1 shows the layout of the files, using fictitious data and variable names for illustrative purposes. The first column shows PIKs, or unique person identifiers; and within a given PIK, each row represents a claim. Any time an SSN is created or information associated with an existing SSN is changed, that event is registered as a claim. Table 1 provides examples of claims associated with changes to given names, surnames, date of birth, and sex-coding. For instance, John Doe (PIK=01) corrected a transposed month and

[5]Therefore, it is important to stress that people who change both their name and their sex-coding did not *necessarily* undergo SRS.

[6]Specifically, 5 U.S.C. §552a (b) (4), 13 U.S.C. §6, and 42 U.S.C §902 and §1306. Ensuring confidentiality of the data are the primary concern at the Census Bureau; data stewardship training is required annually for all staff, and severe penalties exist for any misuse of data.

[7]The Census Bureau is analyzing the use of administrative records to improve data quality and reduce respondent burden in its surveys and censuses. Record linkage in Census Bureau processes or products requires administrative records with consistent and high-quality information on name, date of birth, and sex. Researching the transgender population provides valuable insight to researchers and practitioners by exploring how name and sex-coding changes legitimately appear in the SSA file and other administrative data sources. This is an important step in understanding that changes in first name and sex data are not necessarily anomalies or errors.

day in his date of birth record; Jane Smith (PIK=02) changed her surname to Doe, consistent with the popular convention of married women adopting their husbands' surnames; and John Miller's (PIK=10) apparently mis-entered sex-coding was corrected when he was less than 2 years old.

In order to identify likely transgender people, the first task is to distinguish between people whose first name or sex-coding changed and those whose first name and sex-coding were stable over the observation period. In the example in Table 1, PIKs 01–06 are "stable", while PIKs 07–10 feature a change in the first name, sex-coding, or both.

Table 2 shows the total number of claims, the total number of unique PIKs, and the average number of claims per PIK within the universe of analysis. These figures are also given for the set of PIKs with stable first name and sex-coding information and the set of PIKs featuring a change in first name, sex-coding, or both. After dropping those who were not alive and over 16 years old on 1 April 2010, my data include 658.5 million claims associated with 308.6 million unique PIKs. For those over 16 years old and alive on 1 April 2010, the average person had 2.1 claims associated with their Numident record. For the vast majority of PIKs—290.1 million, or 93.9 percent—we observe no first name or sex-coding changes. Nevertheless, these "stable" records have 2.0 claims associated with their Numident record on average. Changes to these records might include corrections to erroneous dates of birth (as in John Doe's (PIK=01)), changes of surnames (as in Jane Smith's record (PIK=02)), or records of death (as in Jane Johnson's record (PIK=08)). The remaining 18.6 million PIKs exhibit at some point a change in their first names, sex-coding, or both. I next describe how I sort through these 18.6 million PIKs, and their 64.7 million claims, to identify PIKs whose first name and sex-coding changes are consistent with a gender transition.

2.3 Modeling Gender Transitions

To identify those who are likely to have undergone a gender transition, I look first for those who changed their first name from a traditionally male name to a traditionally female name, or vice versa. Among those whose name changes are consistent with gender transition, I then identify people whose sex-coding changed in the same direction as their first name. To simplify my analysis and to reduce the likelihood of false-positive assignment of transgender status, I ignore those who changed their first name or sex-coding before turning 16 years old as well as those who later changed their first name or sex-coding back to the original version. Note that these rules need not apply in principle. People can and do transition before turning 16. Further, gender can be a fluid concept; people can (and do) transition from one gender to another and back again.[8] These rules are instead meant to limit the possibility of falsely identifying a person as transgender.

2.3.1 Determining the Gender of a Name

Using the 377.0 million *stable* records, I construct likelihood tables giving, for every name, the proportion whose sex-coding is "M" and the proportion that is "F". Since the gender of names changes over time, I generate these likelihood tables by birth-decade. I then merge these likelihoods, by name and birth decade, onto the 19.3 million records with *non-stable* first names, first by the original first name and then by the new first name. I can identify name changes that are consistent with gender transitions by comparing the likelihood that the original first name is male to the likelihood the new first name is male.

Within the context of Table 1, PIKs 1–6 are used to construct the likelihood tables. The name John is associated with an "M" sex-coding 100 percent of the time, and the name Jane is associated with an "M" 0 percent of the time. However, the name Val is associated with an "M" only 50 percent of the time.

Figure 1 shows the distribution of likelihoods that a name is male. The sharply bi-modal shape of the distribution makes it immediately clear that most names are either strongly male or strongly female. Ap-

[8] For example, Weitze and Osburg (1996) identified a small number of people who transition from one gender to another and back again in Germany.

proximately 39.0 percent of all names in 377.0 million stable records *always* belonged to people whose sex-coding was "M", while about 52.0 percent of all names belonged to people whose sex-coding was "F". Only about 2.0 percent of all names were perfectly gender-neutral.

I have no prior notion of what likelihood threshold is sufficient to confidently classify a name as likely male or likely female, so I present results for three progressively demanding sets of thresholds: 0.1 and 0.9, 0.05 and 0.95, and 0.01 and 0.99. That is, for the first set, a name is categorized as likely male (female) if 90 percent or greater (10 percent or lower) of all stable records with that name had a sex-coding that was "M" ("F"). I call this the *90 percent confidence threshold*. I similarly define the *95 percent confidence threshold* and the *99 percent confidence threshold*. Note that name combinations that meet the 99 percent threshold also meet the 90 and 95 percent thresholds, so these are not mutually exclusive categories.

This framework considers a name change to be consistent with a gender transition if (a) both the original and the new first name fall within the specified confidence threshold, (b) the gender of the new first name is different than the gender of the original first name, and (c) the name change is not reversed later. To illustrate the application of these rules, we return to Table 1. Jane Johnson's (PIK=08) name change from John is considered consistent with a gender transition, since both John and Jane surpass the 90 percent confidence threshold. The same logic applies to John Brown's (PIK=09) name change from Jane Brown. However, Jane Thompson's (PIK=07) name change from Val is not classified as consistent with a gender transition, since the name Val does not meet the minimum confidence threshold.

Applying these rules provides a more inclusive estimate of the U.S. transgender population, since it does not require changes in sex-coding, which typically require costly and not always preferred genital sexual reassignment surgery. Nevertheless, these inclusive estimates are still likely to be *under*estimates for the several reasons discussed in Sections 1 and 3.1.

2.3.2 Sex-Coding Changes

I also identify the subset of those who exhibited a transgender-consistent name change who also changed the sex-coding associated with their SSNs (and did not later reverse that change). Of those, I further identify those for whom the direction of the sex-coding change matched the direction of the name change. This provides the most conservative estimate of the U.S. transgender population. Returning to the example in Table 1, we see that four years after Jane Johnson (PIK=08) changed her first name from John, she changed her sex-coding from "M" to "F". At the same time that John Brown (PIK=09) changed his name from Jane, he also changed his sex-coding from "F" to "M". John Miller's (PIK=10) gender-maker change, however, is not consistent with a gender transition for two reasons. First, his first name did not change. Second, his sex-coding changed before he turned 16. Such scenarios are relatively common among sex-coding changes, and appear to reflect corrections to mis-entered sex-coding during the process of applying for an SSN, which is why the age-rule is important for our estimates. These estimates are presented below.

3 Results

3.1 Estimates of the Transgender Population

Table 3 shows the estimated number of transgender SSN holders over 16 years old who were alive in the U.S. on 1 April 2010. The top panel gives estimates for the entire population, while the bottom two panels give estimates separately for Male to Female (MTF) and Female to Male (FTM) transgender people. The first row within each panel gives the number of people who changed their first name in a way consistent with gender transition, the second row gives the number who also changed their sex-coding, and the third row gives the number who changed their sex-coding in the same direction as their name change. Thus, movement down the rows represents increasingly restrictive definitions of transgender. Moving from left to right across

columns gives the estimates using the 90 percent, 95 percent, and 99 percent confidence thresholds for the gender of the original and new first names. Since these confidence thresholds are increasingly demanding, the most conservative estimate of the U.S. transgender population appears in the bottom-right cell of each panel, and the most inclusive estimate appears in the top-left cell.

According to the most inclusive criteria, which require a name change consistent with gender transition using the 90 percent confidence threshold, there were 359,274 likely transgender adult SSN holders on Census day 2010. This estimate includes any individuals who changed their first name from a traditionally female name to a traditionally male name (or vice versa), which is consistent with that person taking steps to be socially recognized as a gender distinct from the sex assigned at birth, regardless of whether that person has taken steps to alter his or her physical appearance through hormone treatment or surgery. This estimate of the transgender population, while the most inclusive within the context of my estimation strategy, is still substantially smaller than Gates's (2011) estimate of 697,529. That estimate derives, however, from survey estimates of the transgender populations in California and Massachusetts; if transgender adults disproportionately cluster in those two states—as suggested by Rosser *et al.* (2007)—then that would bias Gates's estimate upward. At the same time, my estimate of 359,274 does not include those who have not legally changed their names, who have changed their names through common-law, whose name(s) do no meet my (admittedly arbitrary) minimum confidence threshold, or whose gender transition did not involve a name change at all. Furthermore, my estimates do not include people who do not possess SSNs, who transitioned before attaining an SSN, or who transitioned before turning 16. For these reasons, my estimates are probably underestimates. It is not surprising that as the name confidence thresholds become more demanding, moving rightward across the columns, the estimates attenuate; the 99 percent confidence threshold yields an estimate of 126,679.

The more demanding estimate requires individuals to alter their sex-coding in the same direction as their name change. As noted above, during 2010 individuals were required to submit written evidence of genital SRS in order to change their sex-coding with the SSA. Since many transgender people cannot or choose not to undergo SRS and nevertheless live fully as the gender with which they identify, these estimates reflect more the subset of transgender people who have had SRS and have successfully navigated the process of changing their documentation to affirm their gender.[9] Insofar as these estimates typically require SRS, they are roughly comparable to the literature that uses medical registries and surveys of medical practitioners to estimate the prevalence of transsexualism (e.g., Bakker *et al.*, 1993; De Cuypere *et al.*, 2007; Tsoi, 1988; Wålinder, 1968; Wilson *et al.*, 1999).

The number of adult SSN holders who changed their first name and sex-coding in the same direction ranges from 17,481 (using the 99 percent confidence threshold) to 24,784 (using the 90 percent confidence threshold). Since there were about 247.5 million U.S. residents age 15 and over in 2010 (Census Bureau, 2013b), about 93 percent of whom were estimated to be citizens (Census Bureau, 2013a), these estimates imply that 0.008 to 0.011 percent (or 7.6 to 10.8 out of 100,000) of adult SSN holders had undergone SRS by 1 April 2010. This estimate compares favorably with that of Wilson *et al.*'s (1999) estimate of 0.008 percent in Scotland, but is substantially larger than Wålinder's (1968) estimate of 0.002 percent of adults in Sweden, O'Gorman's (1982) estimate of 0.002 percent in Ireland, and Pauly's (1968) estimate of 0.001 percent in the U.S.

Turning to the bottom two panels, my estimates suggest that 12,320 to 16,755 adult SSN holders underwent MTF SRS, while 5,161 to 8,029 adult SSN holders underwent FTM SRS. This suggests that about 0.011 to 0.015 percent of people who are assigned male gender at birth undergo SRS, while about 0.004 to 0.007 percent of people assigned female gender at birth undergo SRS. The estimates for MTF SSN holders

[9] Spade (2008) notes that the clerk processing a person's request for a sex-coding change often has significant choice over whether or not to approve the request. On the one hand, some clerks have approved the request without requiring the necessary documentation, on the other hand, some clerks have rejected requests for sex-coding changes even when the proper documentation was supplied.

who underwent SRS are larger than those of Ross *et al.* (1981) (0.004 percent in Australia), De Cuypere *et al.* (2007) (0.007 percent in Belgium), Bakker *et al.* (1993) (0.008 percent in the Netherlands), and estimates of 0.003 percent from the American Psychiatric Association (APA, 2000). They are smaller, however, than estimates by Veale (2008) (0.027 percent in New Zealand) and Tsoi (1988) (0.034 percent in Singapore). Similarly, my estimates for FTM SSN holders who underwent SRS are larger than Ross *et al.* (1981) (0.0006 percent), De Cuypere *et al.* (2007) (0.003 percent), Bakker *et al.* (1993) (0.003 percent), and the APA (2000) (0.001 percent); but my estimates are smaller than Tsoi's (1988) (0.012 percent). They are similar to Veale's (2008) estimate of 0.004 percent.

Consistent with previous literature, I find that the rate of MTF transitions involving SRS is higher than FTM transitions involving SRS (see, for example, APA, 2000; Bakker *et al.*, 1993; Tsoi, 1988; Pauly, 1968; Wålinder, 1968). An important finding, however, is that this pattern is reversed when we apply the more inclusive rule for identifying likely transgender people. When considering name changes consistent with gender transition only, there are 1.78 to 2.70 times more FTM likely transgender adult SSN holders than there are MTF likely transgender adult SSN holders. This reversal could reflect differences in costs of SRS (see Meier and Labuski, 2013, for a discussion of costs of SRS), differences in perceived effectiveness of genital SRS for FTM people versus MTF people, or differences in preferences. What is clear is that the more conservative estimation strategy, which is similar to the common studies evaluating medical registries or censuses of clinics, misses the largest component of the transgender population and masks nuance in the ways people choose to transition.

3.2 Transition Paths

The structure of the SSA update files allows us to investigate the paths people take in documenting their gender transition. Table 3 demonstrates, among other things, that there are many more people who change their first names in ways that are consistent with gender transitions than there are people who change their first names as well as their sex-coding. But of this second group, what paths to do people take to ensure their legal records reflect their asserted gender? At what age do people typically begin to document gender transition? Do name changes and sex-coding changes occur simultaneously, or do people change their names first and their sex-coding later? If they stagger the changes, how much time lapses between changes?

Table 4 helps answer these questions. I present counts of those who changed their name first, counts of those who changed their sex-coding first, and counts of those who changed their names and sex-coding simultaneously. For each group, I also show the mean age at the time of the first change (i.e., name or sex-coding), and the mean number of years between that change and the second change. I present this information for those whose names met the 90 percent confidence threshold but not the 95 percent threshold, those whose names met the 95 percent threshold but not the 99 percent threshold, and those whose names met the 99 percent confidence threshold or above. As in Table 3, I show results for the entire population, the MTF population, and the FTM population.

Table 4 shows that, among those who changed their first names and sex-coding in ways that were consistent with gender-transition, the majority made both changes at the same time. For example, 66.6 percent of those whose original and new names met the 90 percent but not the 95 percent confidence threshold changed both their name and their sex-coding concurrently. As the requirements on names becomes stronger, we see an increase in the percentage of people who concurrently change their name and sex-coding to 76.1 percent (among those whose names meet the 95 percent but not the 99 percent threshold) and then to 78.7 percent (among those whose names clear the 99 percent threshold). These patterns are similar for the likely MTF and likely FTM groups. The average age at which people make simultaneous name and sex-coding changes ranges from about 34.4 to 39.0 years old, depending upon the confidence threshold applied to the names.

The next most common path is to change the first name and then the sex-coding some time later. Overall, approximately 19 to 22 percent of those who change both their names and their sex-coding change their names

first, although it is slightly more common for the FTM group to change their names before their sex-coding than it is for the MTF group.[10] Among those who change their name before their sex-coding, the mean age at which the name change occurs ranges from 32.7 to 36.6, and on average 5.2 to 6.9 years pass before they change their sex-coding.

A small number of people change their sex-coding before changing their first names. While this group comprises only 1.4 to 5.1 percent of those who change both their names and their sex-coding and also clear the 95 percent confidence threshold, 11.6 percent of those who change their names and sex-coding but only clear the 90 percent confidence threshold change their sex-coding first. This group on average makes the first changes to the information in their SSA records when they are 21.8 to 26.1 years old, which is much earlier than either those who change their names first or those who change their names and sex-coding simultaneously. Furthermore, those who change their sex-coding first wait much longer to change their names (10.8 to 15.7 years) than do those people who change their names before their sex-coding. These differences could be due to at least two possible scenarios. In the first case, people may change their names through common-law rule, rather than through formal means, and their sex-coding with the SSA. Much later, these same people might find it beneficial to formalize the name change that occurred much earlier. Alternatively, this group could reflect measurement error in the underlying data. Since the group is relatively small, I leave this for future research and turn next to an examination of how rates of claims that are consistent with gender transitions have changed over time.

3.3 Historical rates of Transgender-Consistent Claims

Finally, we turn to the question of how the annual flows of claims that are consistent with gender transitions have changed during the years since the SSA was established. Figures 2–4 investigate changes in the frequency of claims that are consistent with gender transitions relative to all other claims. In the discussion that follows, I use data from all records, including those of people who were not alive during the 2010 Census. Transgender-consistent claims are defined as any name changes that meet the 90 percent confidence threshold, and any sex-coding changes that move in the same direction as the name changes. If the name change and sex-coding change occurred separately, then each claim is counted; if the name and sex-coding changes occurred simultaneously, then that is counted as a single claim. All other claims are considered non-transgender-consistent claims and are used as a benchmark of growth in claims due to changes in the population or the legislation and uses of the SSN over time.[11]

We begin by looking at annual counts of claims that are consistent with gender-transitions and counts of all other claims. Figure 2 plots the counts of transgender-consistent claims and all other claims by year, from 1936 to 2010. Transgender-consistent claims are shown with the dashed line and measured on the left-hand axis, while all other non-transgender-consistent claims are shown with the solid line and measured on the right-hand axis. Immediately evident is that there is evidence of gender-transition consistent claims as early as the 1930s. In addition, there is marked growth in both the number of transgender-consistent claims as well as the number of all other claims processed each year. For non-transgender-consistent claims, this growth appears to begin in the late 1940s and continue steadily before accelerating in the early 1970s.

[10] The exception to this is among the group whose names clear the 99 percent confidence threshold.

[11] In 1998, an anomalous spike appeared in the trends of both the non-transgender-consistent claims as well as the transgender-consistent claims. I investigate this and conclude that the spike appeared to be an artifact of the administrative record keeping process, rather than the result of an historical event. Specifically, 1998 was the year that the SSA supplied the U.S. Census Bureau with the 100% Numident file, containing every transaction for every SSN, from 1936 onward (USCB-SSA, 2000). In most other years, the share of claims is evenly distributed across months, however this is not the case in 1998. An unusually large number of claims were filed in January. I suspect that a relatively small number of transactions from 1936 onward were missing claim entry dates, and that these missing dates were replaced with the date 1 January of 1998. In order to better display the trends over the rest of the years, I impute values for 1998 by averaging the values in 1997 and 1999. Results using the data prior to the imputation, as well as the results of my inspection of 1998, are available upon request.

On the other hand, growth in transgender-consistent claims does not appear to pick up until the 1970s. This decade was a time of rapid expansion of the use of the SSN for purposes outside the Social Security Program. For example, new legislation required financial institutions to obtain SSNs of all their customers in 1970; in 1975 all recipients of federal benefits were required to provide an SSN; and 1976 brought legislation that allowed states to require SSNs for taxes, eligibility for state programs, and motor vehicle registrations (Puckett, 2009). In addition, the Social Security Administration began to store information on each SSN in an electronic Numeric Identification System ("the Numident") in 1972, and the next several years were spent digitizing previous paper records, all of which could contribute to the sharp rise in claims during the 1970s. These factors seemed to have a similar impact on the number of transgender-consistent and non-transgender-consistent claims during that time. By 1979, the number of claims had dropped back to its 1960s levels. The sharp spike in non-transgender-consistent claims in 1987 is likely attributable to the Immigration Reform and Control Act, which was signed into law toward the end of the previous year and resulted in approximately 3 million immigrants gaining legal status (Bean *et al.*, 1989). In 1987, the SSA also introduced Enumeration at Birth (EAB), which allows parents of newborns to register their children for an SSN as part of the birth registration process (Puckett, 2009; SSA, 2010). Interestingly, transgender-consistent claims do not exhibit the same spike in 1987, although a sharp increase in transgender-consistent claims occurred in 1988. During the remainder of the 1990s and early 2000s, the total number of non-transgender-consistent claims per year stayed relatively constant, with a small spike in 1993.[12] Transgender-consistent claims, on the other hand, showed a steady decrease during the 1990s and 2000s, with a brief jump in 2000. While the difference in scale makes it difficult to identify changes that are unique to the transgender population versus the broader population, these annual counts show clearly the growth in both types of claims.

To help identify years in which growth in transgender-consistent deviated from the rate of growth in other types of claims, we next consider Figure 3, which plots the annual percent change in transgender-consistent claims, de-trended by the annual percent change in non-transgender consistent claims. That is, Figure 3 shows the difference between the year-over-year growth rate of transgender-consistent claims and non-transgender-consistent claims. We can see that after the first five years of the SSA and up until 1975, transgender-consistent claims grew at a pace that was about 10 percentage points lower than the growth in all other claims. The expansions in the use of the SSN in the 1970s appear to have had larger impact on the growth of transgender-consistent claims than non-transgender-consistent claims. The mid-1970s saw some of the first political gains for transgender constituents, including the Connecticut District Court's decision that a transgender person had the right to have the sex-coding changed on her birth certificate; the passing of the first city ordinance protecting the civil rights of transgender people in Minneapolis; and the New York State Supreme Court decision that a transgender woman, Renée Richards, was eligible to play in the U.S. Open tennis championships (Darnell v. Lloyd, 1975; City of Minneapolis, 1975; Amdur, 1977). It is possible that those events encouraged the especially high growth in transgender-consistent claims during those years and, to a lesser extent, the decade that followed. The most evident difference in claim growth rates, however, is in 1988. In that year, the growth rate of transgender-consistent claims was over 200 percentage points greater than the growth rate of other claims. I did not identify major events in U.S. transgender history or in the history of the SSA that could explain this feature of the data, so I must leave the question of its cause to future research. Finally, a substantial jump in the de-trended growth rate of transgender-consistent claims is observed in 2000, followed by a deep drop in 2001–2003, which corresponds to the introduction of stricter rules regulating sex-coding changes (Spade, 2008). At the same time, increased concern over security during the onset of the War on Terror led the Department of Homeland Security (DHS) and the SSA to begin sending letters to employers whose employees' stated information on name and sex did not match

[12] This was possibly due to individuals filing for Social Security Benefits in anticipation of the effects of the Omnibus Budget Reconciliation Act, signed into law on 10 August 1993, which increased the potential tax liability of high-income Social Security beneficiaries (DeWitt, 2001).

the records associated with their SSNs. As a result, several people were outed as transgender at work, which may have increased the incentive for transgender people to change the sex-coding on their files.[13] To help disentangle the influence of these events, I next present annual counts of each *type* of transgender-consistent claim: first name changes, sex-coding changes, and simultaneous changes.

Figure 4 shows the annual counts of first name changes (shown with the dashed line and measured), sex-coding changes (shown with the solid line), and simultaneous changes (shown with the dotted line). First name changes are measured on the left-hand axis, while sex-coding and simultaneous changes are measured on the right-hand axis. Figure 4 demonstrates that, while the different types of transgender-consististent often follow the same patterns of growth and decline, there are also instances in which the patterns vary across the different types of claims. For example, all types of transgender-consistent claims occur, but are rare, until they become more common in the late 1960s and especially the 1970s. Counts of each type of claim exhibit a local peak in 1976. However, 1982 shows a large jump in simultaneous (sex-coding and first name change) claims, while separate claims for first name changes and sex-coding changes do not have nearly as dramatic an increase in that year. By 1988, though, all three types of claims increased substantially. Finally, in 2001 and through the rest of the decade, we see a drop in name changes accompanied by a growth in sex-coding changes and a leveling-out in simultaneous changes. This pattern is consistent with individuals responding to the incentives of the War on Terror policies by reducing the number of name changes unaccompanied by sex-coding changes and increasing the number of sex-coding changes (among those who had already changed their names).

4 Conclusion

This paper presents estimates of the U.S. adult transgender population on 1 April 2010 using administrative records from the Social Security Administration. Two main estimates are given. The first identifies all adult SSN holders who changed their first name from a high-probability male name to a high-probability female name (or vice versa) *and* changed the sex-coding associated with their SSN in the corresponding direction.

The second estimate drops the requirement that the sex-coding was changed, and only requires the person to have changed his or her name in a way that is consistent with a gender transition. Because many transgender people do not wish to undergo SRS or cannot afford it, this more inclusive estimate is seen as corresponding more closely to the number of people who have transitioned in their social and professional lives, regardless of what treatment they may or may not have undergone.

I find that the transgender population is substantially larger than what we would expect if we only looked at estimates of the population that changed both their name and their sex-coding. While my estimates suggest that there were 17,481–24,784 adult SSN holders who changed both their name and their sex-coding in a way that is consistent with a gender transition, I find that 126,679–359,274 adult SSN holders had changed their names in ways that are consistent with gender transitions. I also find that, among those who change both their names and their sex-coding, the majority change both characteristics on their SSA records at the same time; however approximately 20 percent change their name first and then their sex-coding 5–6 years later.

My analysis of the flows of transgender-consistent claims within the SSA data yield several interesting findings. First, as early as the 1930s, people were changing the information associated with their SSN in ways that were consistent with gender transitions. Yet in the 1970s, a period of increasing use of the SSN as a unique identifier as well as a period that included several early advances in the recognition of transgender people's rights, transgender-consistent claims grew at a faster rate than all other types of claims. Finally, in the early 2000s, the number of transgender-consistent name change claims dropped rapidly, while the

[13] The SSA ended this policy in September, 2011 (NCTE, 2011).

number of sex-coding change claims increased, possibly as a result of the increased scrutiny of people's identity documentation during the War on Terror.

This research design has several strengths, but it also comes with a host of limitations. Specifically, there are several groups of transgender people that my estimation strategy necessarily fails to identify. These groups include people who do not have a Social Security Number; people who have an SSN but transitioned before obtaining that SSN or transitioned before turning 16 years old; people whose original or new first name could not be assigned a gender with strong-enough confidence; people who have changed their names through common law and did not register the name change with the SSA; and people who were unable to overcome the substantial financial, temporal, institutional, and administrative barriers to legally changing their names. Finally, all those who experience gender dysphoria but have not taken steps to transition in daily life are not included in my estimates. On the other hand, some intersex people who have rejected the sex assigned to them by parents and physicians shortly after birth may be wrongly classified as transgender. Because this last group is likely to be much smaller than the several groups listed above, my estimates should be interpreted as probable *under*estimates.

Despite these limitations, this research contributes to the existing literature in many ways. First, the methodology used in this paper provides a powerful alternative to the more common approach of looking at medical registers or surveying physicians and other specialists in the field of gender dysphoria. Nevertheless, my more conservative estimates using name changes and corresponding sex-coding changes allow me to compare my results to results deriving from those more common methodologies. At the same time, my approach allows me to identify a much broader portion of the transgender population by looking only at name changes and dropping the sex-coding change requirement. This research also demonstrates the way that administrative records can be utilized to answer questions that are difficult, impractical, or impossible to answer using conventional survey or census data.

My results provide major empirical contributions by updating 45 year old estimates from Pauly (1968). Furthermore, my estimates are likely more representative of the entire U.S. population, since they do not rely on the assumption that transgender people dwell evenly across all states, as in the recent brief by Gates (2011). My analysis of the sequence of events documenting gender transition gives a unique look into the process of transitioning. And my treatment of the historical trends in transgender-related claims sheds new light on the cultural and legal history of transgender people in the United States.

This paper is the first step in a broader research agenda on the social and economic status of transgender people in the United States. Record linkage work could allow for future work that would match likely transgender people identified in this paper to their records in the decennial Census and several years of the American Community Survey. A primary objective of research using such linked data will be to estimate differences between transgender and non-transgender people (if any) in labor coding outcomes such as income and employment. However, such linked data will also permit us to look at differences in access to health care, settlement patterns, and housing. Further analysis of the historic trends could also allow us to investigate whether changes in policies governing the documentation of gender affected rates of transgender-consistent claims with the SSA.

References

Amdur, N. (1977) Renée richards ruled eligible for u.s. open, *New York Times, August 17 1977*, p. 40.

American Psychiatric Association (APA) (2000) *Diagnostic and statistical manual of mental disorders: DSM-IV-TR®*, American Psychiatric Pub.

Bakker, A., Kesteren, P. v., Gooren, L. and Bezemer, P. (1993) The prevalence of transsexualism in the Netherlands, *Acta Psychiatrica Scandinavica*, **87**, 237–238.

Bean, F. D., Vernez, G. and Keely, C. B. (1989) *Opening and closing the doors: Evaluating immigration reform and control*, vol. 1, The Urban Insitute.

City of Minneapolis (1975) Minneapolis, Minnesota, Code of Ordinances, Title 7 Civil Rights, Chapter 139, `http://library.municode.com/HTML/11490/level2/COOR_TIT7CIRI.html`.

Conron, K. J., Scott, G., Stowell, G. S. and Landers, S. J. (2012) Transgender health in Massachusetts: results from a household probability sample of adults, *American Journal of Public Health*, **102**.

Darnell v. Lloyd (1975) 395 F. Supp. 1210, *Civ. No. H-75-33*.

De Cuypere, G., Van Hemelrijck, M., Michel, A., Carael, B., Heylens, G., Rubens, R., Hoebeke, P. and Monstrey, S. (2007) Prevalence and demography of transsexualism in Belgium, *European Psychiatry*, **22**, 137–141.

DeWitt, L. (2001) Research Note # 12: Taxation of Social Security Benefits, *Research Notes and Special Studies by the Historian's Office*.

Dvorak, P. (2012) Transgender at Five, *The Washington Post, May 20 2012*, `http://articles.washingtonpost.com/2012-05-19/local/35457768_1_jean-kathryn-transgender-people`.

Gates, G. J. (2011) How many people are lesbian, gay, bisexual and transgender?, *The Williams Institute*.

Grant, J. M., Mottet, L. A., Tanis, J., Harrison, J., Herman, J. and Keisling, M. (2011) Injustice at Every Turn: A Report of the National Transgender Discrimination Survey, *Washington: National Center for Transgender Eaquality and National Gay and Lesbian Task Force*.

Landén, M., Wålinder, J. and Lundström, B. (1996) Incidence and sex ratio of transsexualism in Sweden, *Acta Psychiatrica Scandinavica*, **93**, 261–263.

Lovett, I. (2013) California: Rights Guaranteed for Transgender Students, *The New York Times, August 12 2013*, `http://www.nytimes.com/2013/08/13/us/california-rights-guaranteed-for-transgender-students.html?smid=pl-share`.

Lu, A. (2011) For Transgender Detainees, a Jail Policy Offers Some Security, *The New York Times, December 23 2011*, `http://www.nytimes.com/2011/12/23/us/for-transgender-detainees-a-jail-policy-offers-some-security.html`.

Meier, S. C. and Labuski, C. M. (2013) The demographics of the transgender population, in *International Handbook on the Demography of Sexuality*, Springer, pp. 289–327.

National Center for Transgender Equality (NCTE) (2011) Social Security Administration No Longer Notifying Employees About Gender, *Press Release, September 15 2011*, `http://transequality.org/news11.html`.

O'Gorman, E. C. (1982) A retrospective study of epidemiological and clinical aspects of 28 transsexual patients, *Archives of Sexual Behavior*, **11**, 231–236.

Pauly, I. B. (1968) The current status of the change of sex operation, *The Journal of Nervous and Mental Disease*, **147**, 460–471.

Puckett, C. (2009) The story of the Social Security Number, *Social Security Bulletin*, **69**, 55–74.

Reed, B., Rhodes, S., Schofield, P. and Wylie, K. (2009) Gender variance in the UK: Prevalence, incidence, growth and geographic distribution, *Gender Identity Research and Education Society*.

Ross, M. W., Wålinder, J., Lundström, B. and Thuwe, I. (1981) Cross-cultural approaches to transsexualism, *Acta Psychiatrica Scandinavica*, **63**, 75–82.

Rosser, B. S., Oakes, J. M., Bockting, W. O. and Miner, M. (2007) Capturing the social demographics of hidden sexual minorities: An internet study of the transgender population in the United States, *Sexuality Research & Social Policy*, **4**, 50–64.

Ryan, H. (2013) TV's Transformative Moment: Netflix's New Show Takes a Casting Cue from Real Life, *Newsweek, July 17 2013*, **161**.

Social Security Administration (SSA) (2010) Program Operations Manual System, RM 00202.308 Enumeration at Birth Process, `https://secure.ssa.gov/poms.nsf/0/4b2e34a39e55f770852576b30009b02a!OpenDocument&Click=/`.

Social Security Administration (SSA) (2013) Program Operations Manual System, RM 10212.200 Changing Numident Data for Reasons other than Name Change, `https://secure.ssa.gov/poms.nsf/lnx/0110212200`.

Sosin, K. (2013) New study to weigh transgender military inclusion, *Windy City Times, August 4 2013*, `http://www.windycitymediagroup.com/lgbt/New-study-to-weigh-transgender-military-inclusion/43930.html`.

Spade, D. (2008) Documenting gender, *Hastings Law Journal*, **59**, 731–842.

Spiegel, A. (2008) Two Families Grapple with Sons' Gender Identity: Psychologists Take Radically Different Approaches In Therapy, *National Public Radio, May 07 2008*, `http://www.npr.org/2008/05/07/90247842/two-families-grapple-with-sons-gender-preferences`.

Stanglin, D. (2013) Navy SEAL's biggest secret: A transgender life, *USA Today, June 5 2013*, p. 3A.

Tsoi, W. F. (1988) The prevalence of transsexualism in Singapore, *Acta Psychiatrica Scandinavica*, **78**, 501–504.

U.S. Census Bureau and Social Security Administration (USCB-SSA) (2000) 2000 Memorandum of Agreement Between the Bureau of the Census and the Social Security Administration Numident File, *Internal Document*.

U.S. Census Bureau (Census Bureau) (2013a) American Community Survey 2010, American Community Survey 1-Year Estimates, Table B05001, generated using American Fact Finder, `http://factfinder2.census.gov/bkmk/table/1.0/en/ACS/10_1YR/B05001/0100000US`.

U.S. Census Bureau (Census Bureau) (2013b) Census 2010, Summary File 1, Table DP-1, generated using American Fact Finder, http://factfinder2.census.gov/bkmk/table/1.0/en/DEC/10_DP/DPDP1/0100000US.

Veale, J. F. (2008) Prevalence of transsexualism among New Zealand passport holders, *Australasian Psychiatry*, **42**, 887–889.

Votta, R. (2012) 'Glee' Recap: Disco Nightmare? Try Fun Diversion, *Billboard.com*, http://www.billboard.com/articles/news/493448/glee-recap-disco-nightmare-try-fun-diversion.

Wålinder, J. (1968) Transsexualism: definition, prevalence, sex distribution, *Acta Psychiatrica Scandinavica*, **43**, 255–258.

Weitze, C. and Osburg, S. (1996) Transsexualism in Germany: empirical data on epidemiology and application of the German Transsexuals' Act during its first ten years, *Archives of Sexual Behavior*, **25**, 409–425.

Wilson, P., Sharp, C. and Carr, S. (1999) The prevalence of gender dysphoria in Scotland: a primary care study., *The British Journal of General Practice*, **49**, 991.

Appendix 1 Tables

Table 1: Sample Record Layout of Social Security Updates

PIK	FIRST NAME (FN)	LAST NAME (LN)	SEX	ENTRY DATE	BIRTH DATE	DEATH DATE	CLAIM TYPE
01	JOHN	DOE	M	19650101	19640501		CREATION
01	JONN	DOE	M	19800501	19640105		DOB CORRECTION
02	JANE	SMITH	F	19660224	19660214		CREATION
02	JANE	DOE	F	19870715	19660214		LN
03	VAL	JONES	F	19670616	19670601		CREATION
04	VAL	WILLIAMS	M	19601030	19600909		CREATION
05	JOHN	GREEN	M	19621015	19620909		CREATION
06	JANE	WHITE	F	19640215	19631102		CREATION
07	VAL	THOMPSON	M	19680720	19680704		CREATION
07	JANE	THOMPSON	F	20071215	19680704		FN
08	JOHN	JOHNSON	M	19640307	19640207		CREATION
08	JANE	JOHNSON	M	19980101	19640207		FN
08	JANE	JOHNSON	F	20020101	19640207		SEX-MARKER
08	JANE	JOHNSON	F	20090801	19640207	20090601	DEATH RECORD
09	JANE	BROWN	F	19700315	19690815		CREATION
09	JOHN	BROWN	M	20050815	19690815		FN AND SEX-MARKER
10	JOHN	MILLER	F	19650320	19650220		CREATION
10	JOHN	MILLER	M	19660501	19650220		SEX-MARKER

Notes: Data and variable names are fictitious and are used for illustrative purpose only.

Table 2: Number of Claims, Number or Unique Records, and Mean Claim per Record

	Number of Claims	Number of Unique Records	Claims per Record
Total	658,460,272	308,640,732	2.1
Stable Records	593,711,674	290,077,055	2.0
Records with name or sex-coding change	64,748,598	18,563,677	3.5

Source: SSA Numident, 1936–2010

Notes: Individuals must be alive and at least 16 years old on 1 April 2010.

Table 3: Estimates of US Transgender Population alive on 1 April 2010

	Confidence Threshold:		
	90 percent	95 percent	99 percent
Panel A. Total			
Name change	359,274	245,828	126,679
Name and sex-coding change	25,308	23,369	17,696
Consistent name and sex-coding change	24,784	22,980	17,481
Panel B. MTF			
Name change	97,229	76,331	45,560
Name and sex-coding change	16,755	15,663	12,320
Consistent name and sex-coding change	16,755	15,663	12,320
Panel C. FTM			
Name change	262,045	169,497	81,119
Name and sex-coding change	8,553	7,706	5,376
Consistent name and sex-coding change	8,029	7,317	5,161

Source: SSA Numident, 1936–2010

Notes: Individuals must be alive on 1 April 2010, at least 16 years old at the point of the name change or sex change, and both changes must not be reversed.

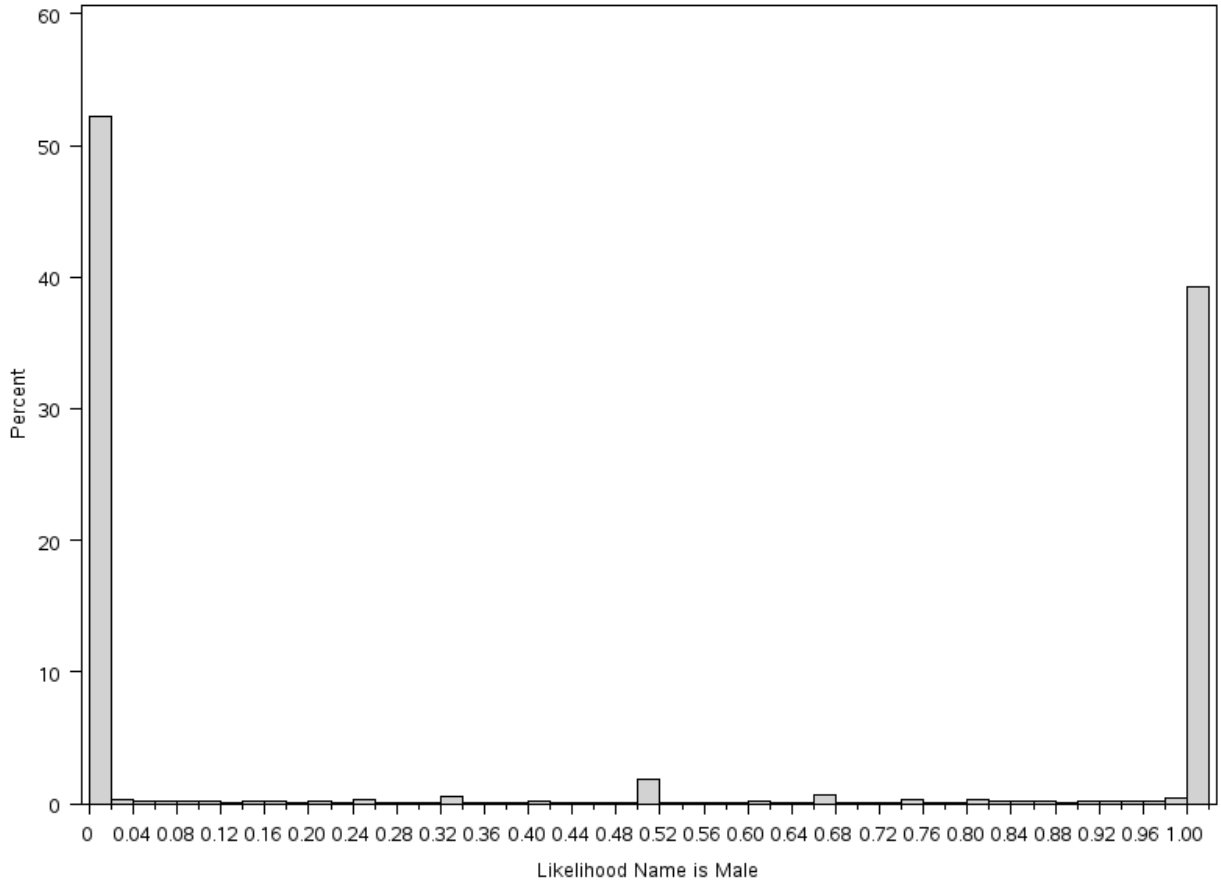
Table 4: Sequences of Transition Among Name and sex-coding Changers

	Confidence Threshold											
	90 percent up to 95 percent				95 percent up to 99 percent				99 percent and above			
	Count	Percent of sample	Age at first change	Years between changes	Count	Percent of sample	Age at first change	Years between changes	Count	Percent of sample	Age at first change	Years between changes
Panel A. Total												
Name then sex-coding	392	0.22	33.1	6.9	1,033	0.19	32.7	5.6	3,483	0.20	36.6	5.2
Sex-coding then Name	210	0.12	21.8	15.7	280	0.05	24.6	12.8	247	0.01	26.1	10.8
Sex-coding and Name	1,202	0.67	34.4	0.0	4,186	0.76	39.0	0.0	13,751	0.79	34.9	0.0
Panel B. MTF												
Name then sex-coding	216	0.20	34.0	6.5	564	0.17	34.9	5.7	2,611	0.21	38.3	5.0
Sex-coding then Name	132	0.12	20.3	16.3	195	0.06	25.2	13.0	184	0.01	26.0	10.8
Sex-coding and Name	744	0.69	34.2	0.0	2,584	0.77	39.2	0.0	9,525	0.77	34.9	0.0
Panel C. FTM												
Name then sex-coding	176	0.25	32.0	7.3	469	0.22	30.1	5.8	872	0.17	31.6	5.8
Sex-coding then Name	78	0.11	24.4	14.7	85	0.04	23.0	12.5	63	0.01	26.5	10.9
Sex-coding and Name	458	0.64	34.6	0.0	1,602	0.74	38.8	0.0	4,226	0.82	34.9	0.0

Source: SSA Numident, 1936–2010

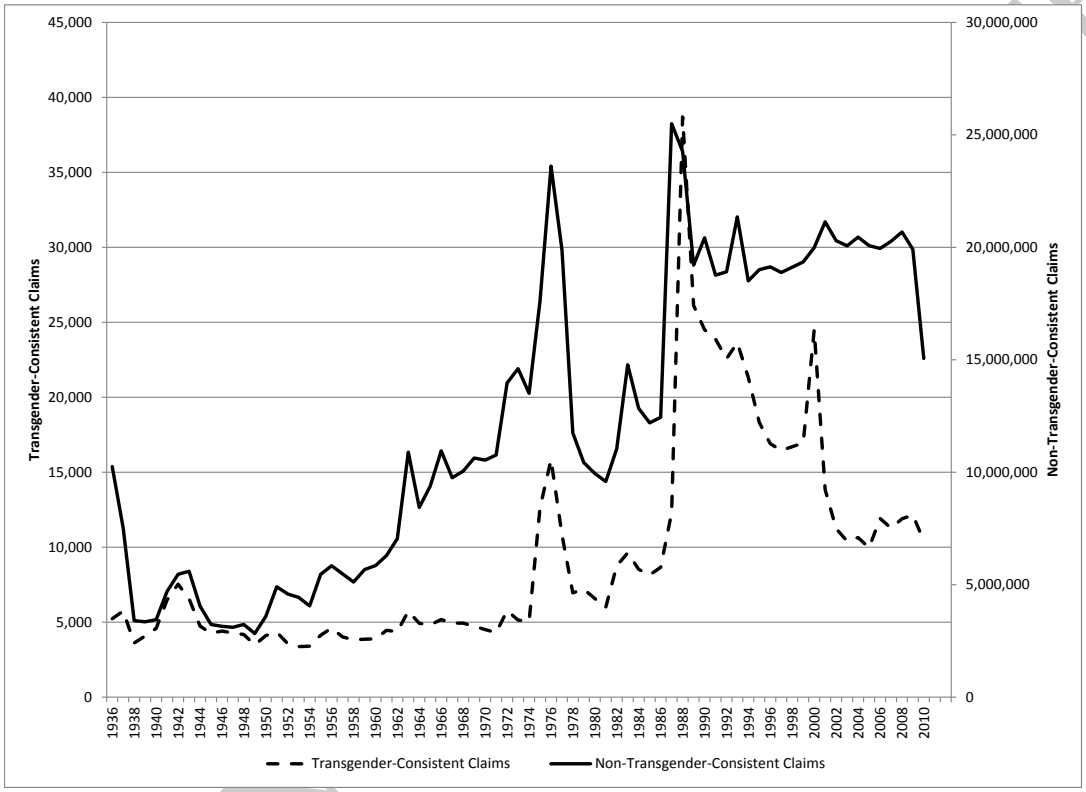
Notes: Figures are calculated over those who changed both their name and their sex-coding in a consistent way. Individuals must be alive on 1 April 2010, at least 16 years old at the point of the name change or sex change, and both changes must not be reversed.

Appendix 2 Figures



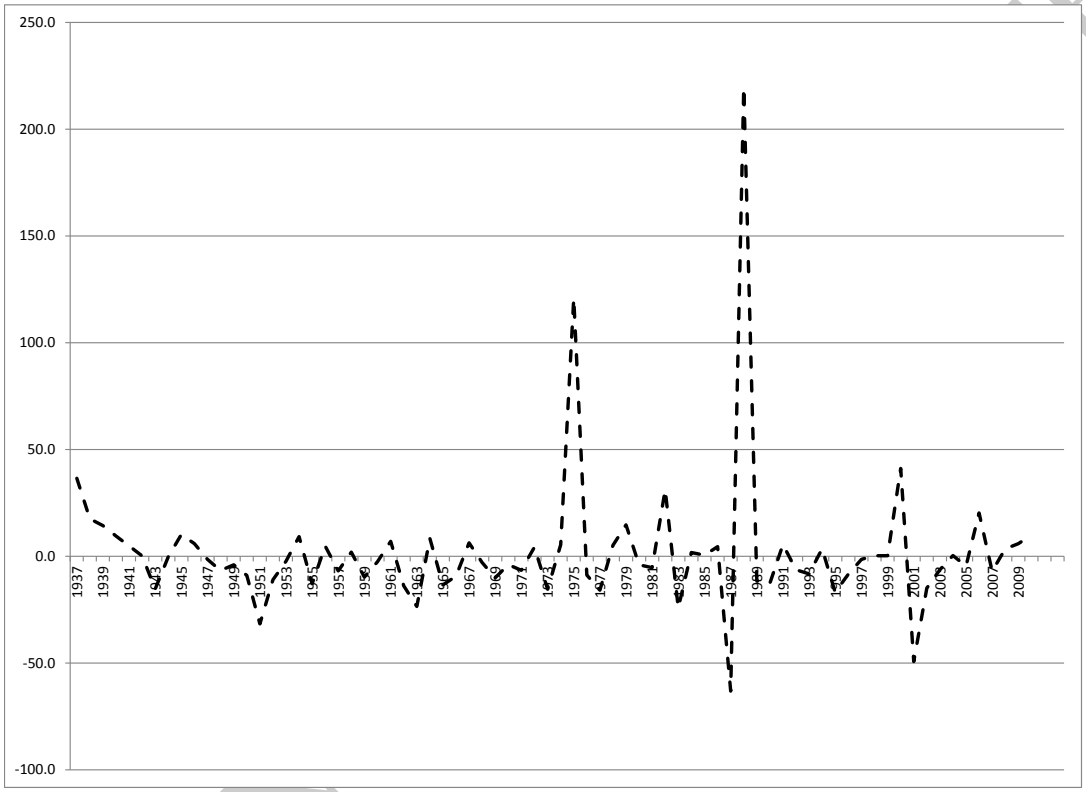
Source: SSA Numident, 1936-2010

Figure 1: Distribution of Likelihoods that a Name is Male



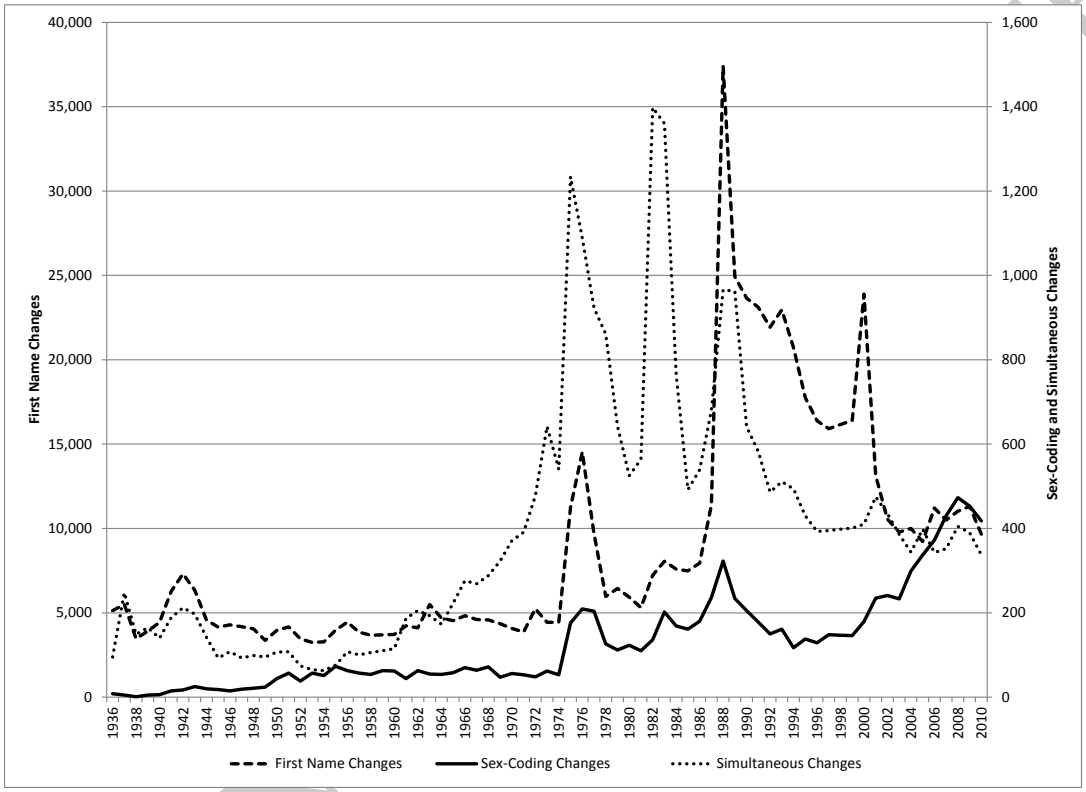
Source: SSA Numident, 1936-2010

Figure 2: Frequency of Transgender-Consistent and All Other Claims



Source: SSA Numident, 1936-2010
Note: Annual percent change in transgender-consistent claims is de-trended by the annual percent change in all other claims

Figure 3: De-trended Annual Percent Change in Transgender-Consistent Claims



Source: SSA Numident, 1936-2010

Figure 4: Frequency of Likely Transgender-Related Claims, by Claim Type