4-30-2021

# Identity for Development: India's Biometric ID Program and Access to Public and Private Services

Meher Kaur
*University of Richmond*

**Identity for Development:**
**India's Biometric ID Program and Access to Public and Private Services**

by

Meher Kaur

Honors Thesis

Submitted to:

Economics Department
University of Richmond
Richmond, VA

April 30, 2021

Advisor: Dr. Alexander Persaud

**Abstract**

Since its rollout in 2009, Aadhaar, India's biometric identification system, has generated unique IDs for nearly 95% of India's population. Since 2009, a variety of public and private services in India have been linked to Aadhaar. This linkage is intended to improve access to public welfare and private services for marginalized groups, such as third gender individuals and rurally located groups, who were historically left out of public welfare programs due to the absence of formal identification. However, there is a general lack of research on how effective Aadhaar is in linking individuals to state and private services and how Aadhaar impacts the lives of marginalized groups. This study employs data from the *State of Aadhaar 2019* report to analyze the effect of Aadhaar attainment on access to two key public services and two key private services: 1) the Public Distribution System (PDS), the National Rural Employment Guarantee Act (MGNREGA), 3) savings accounts, and 4) SIM cards. Data was also used to control for demographic factors such as gender, caste, and rural locality in order to examine how different groups are able to access the select public and private services. In the preferred econometric model, results suggest that Aadhaar has a positive relationship with each of the select services; individuals with Aadhaar have a high probability of using the select public or private service as compared to individuals without Aadhaar. Results also suggest that marginalized groups are less likely to use the select services than their more privileged counterparts, and access to Aadhaar may do little to address this issue.

## Table of Contents

## 1. Introduction

Formal identification is a prerequisite for development in the modern world. The inability to authenticate oneself when interacting with the state—or with private entities such as banks—inhibits access to basic rights and services, including education, formal employment, financial services, voting, social transfers, and more (Gelb, 2013). In this context, biometric identification is playing an increasingly prominent role in today's digital age. Biometric identification uses information such as fingerprints, iris scans, and facial scans to identify individuals. Aadhaar, India's national ID program and the largest biometric ID program in the world, is one of the most ambitious identification programs. It operates on a massive scale and has achieved an astounding level of coverage for a voluntary program.

Aadhaar uses biometric information to allocate a unique identification number to every Indian resident. In 2009, Aadhaar replaced an old, decentralized system of birth certificates and ration cards, which were vulnerable to loss and damage and left marginalized groups struggling to obtain public and private services. Since its rollout in January 2009, Aadhaar has generated biometric identities for over 1.2 billion people across the country, and a variety of services and welfare programs have been linked to Aadhaar IDs (Totapally, 2019).

This linkage to services and welfare provides Aadhaar the ability to eliminate barriers preventing individuals from accessing public and private services in India. For example, historically marginalized individuals in rural India may lack official documentation and transportation to the nearest town, two significant barriers that would prevent them from opening bank accounts. With Aadhaar, individuals can provide verifiable ID and speed up the know-your-customer process that many financial institutions require, thus opening a bank account with relative ease. They can also purchase a SIM card with Aadhaar IDs and access digital financial services without making long trips to a bank. In this way, by linking marginalized groups to state infrastructure and essential private services, Aadhaar can be leveraged as a tool for social inclusion.

Yet, although one of Aadhaar's objectives is to ensure inclusion and social protection for all of India, Aadhaar's impact on minority groups is largely understudied. Household-level data on Aadhaar's implementation is scant and further analysis of economic and sociological impacts of Aadhaar are limited. This paper aims to contribute to the body of literature that evaluates Aadhaar by utilizing existing data on Aadhaar enrollment and usage.

## 1. Research Question and Implications

The goal of this study is to examine the effect of Aadhaar on individuals' ability to access certain key public and private services in India. Enrollment in Aadhaar is expected to positively affect an individual's access to public welfare and private services by eliminating barriers to access and ensuring equitable access. Examples of some of the most important welfare programs in India (that have been linked to Aadhaar) include the Public Distribution System (PDS), the pension payment system, and the National Rural Employment Guarantee Act (MGNREGA). These public programs, in addition to other subsidies and welfare schemes, are vital to the livelihoods of millions of people across India. Relevant private services that are linked to Aadhaar IDs include bank accounts and SIM cards.

In this paper, in order to evaluate Aadhaar's impact on public and private services, I select two of the most important public services and two of the most important private services. Here, "importance" is weighed on the basis of scale and impact. The two key public programs and two key private services are: 1) The Public Distribution System (PDS), 2) the National Rural

Employment Guarantee Act (MGNREGA), 3) bank accounts, and 4) SIM cards. Further explanation of these four services and why they were selected is included in Section 3.

I hypothesize that enrollment in Aadhaar will have a positive effect on access to these four indicators. However, in terms of access to public and private services, I further hypothesize that Indian minorities and marginalized groups will be less likely to obtain access than those that are more economically and socially advantaged.

In the years since Aadhaar's implementation, few public and private bodies have dedicated time and resources to conducting extensive studies of Aadhaar's implementation and effectiveness as a program. I aim to contribute to the analysis of Aadhaar as a program by using pre-existing data to examine the effectiveness of Aadhaar's implementation in linking individuals to state and private infrastructure. These findings have implications for the improvement of Aadhaar as a program. Although Aadhaar has covered 95% of the Indian population, the next step is equitable coverage - ensuring that every individual has access to Aadhaar and is enabled to take advantage of the services it provides. Furthermore, this paper also contributes to the larger body of academic scholarship, such as Gelb (2013) and Khera (2017), that studies the role of identification in international development. As biometric identification continues to be implemented around the world, it is important to examine the social and political implications of this technology and how it can be made even more efficient.

## 2. Background and Literature Review

### 3.1 An Overview of Identification Systems

Kelkar et al (2014) exemplify how birth certificates are often the first way to acquire official identification. The absence of a birth certificate results in difficulty in accessing basic public services, locking individuals in a cycle of marginalization and vulnerability and rendering them as invisible to the state. According to study by UNICEF, in 2000 approximately 36 percent of children worldwide and 40 percent of children in the developing world were not registered at birth. South Asia had the highest percentage of unregistered births, 63 percent. Among the least-developed countries, under-registration in South Asia was 71 percent (UNICEF, 2005) Even for those that are registered, birth certificates are often difficult to access due to poor record keeping, lack of mobility or corruption (Gelb and Clark, 2013) These statistics reveal the need for and importance of a national identification system in India.

Simon Szreter argues that one of the fundamental determinants of an individual's well-being is whether or not the individual has an acknowledged existence (Szreter, 2007). According to Szreter, the right to identity registration at birth is a human right, which almost every state in the world has endorsed through ratification of the United Nations Convention on the Rights of the Child. However, many of the world's poorest countries have no identity or vital registration system. As a result, the poor of the world are anonymous. This impedes their ability to exercise and enjoy their human rights and their entitlements (Szreter, 2007).

Alan Gelb and Julia Clark (2013) reinforce this idea through their working paper titled *Identification for Development: The Biometrics Revolution*, which surveys 160 cases where biometric identification has been used for economic, political, and social purposes in developing countries. The paper raises several significant concerns regarding the implementation of biometric identification. The first issue is of privacy and data protection. Critics of biometrics point to Western notions of securitization and policing that have become inextricably linked to biometric identification (Gelb and Clark, 2013). These points are important to consider when examining the implementation of a massive identification program such as Aadhaar, and similar debates take place within India with regards to Aadhaar as well. The second concern is exclusion. Not every individual is physically able to provide biometrics. Error rates with technology can also lead to exclusion, and careful attention needs to be paid to on-the-ground implementation to ensure that workers operating the authentication technology are aware of the requirements for authentication. Lastly, the third concern is the high cost of biometric technology, and the relatively understudied returns gained from the technology. Without an adequate measure of opportunity costs, it can become difficult to determine true effectiveness of biometric systems. These three issues are also prominent factors in existing academia on Aadhaar.

Gelb and Clark (2013) also highlight several benefits derived from biometric identification, including inclusion and leapfrogging in fragile states. Elimination of fraud is a means of inclusivity since it redirects resources towards the intended beneficiaries. Additionally, with individualized IDs, more individuals can autonomously partake in state and private service distribution systems. Another positive outcome is biometric identification's linkage to a variety of sectors and the generation of positive externalities in each sector (Gelb and Clark, 2013). Several examples include health, elections, and civil service reform. This working paper serves as an important foundation for my research. Many of the points raised by Gelb and Clark (2013) are reflected in prior research on Aadhaar and the findings of this working paper were used in the construction of my econometric model. The plus sides and pitfalls raised provide an important framework for my analysis of Aadhaar's implementation and impact in India.

### 3.2 Overview of Aadhaar

In January 2009, the Unique Identification Authority of India (UIDAI) was constituted by the Indian government as the official body in charge of the creation of a database for the Unique Identification (UID) project and the implementation of the project in India. The idea for this unique ID, later to be named “Aadhaar”, was created by Indian businessman Nandan Nilekani. Nilekani envisioned a more inclusive and just India through the Aadhaar program (Das, 2015). The Indian government’s objectives for creating and implementing Aadhaar are generally twofold: firstly, the program aimed to significantly reduce identity related fraud, reduce leakages, and allow for better targeting of government schemes. Secondly, the program is meant to enhance social inclusion by officially identifying every one of India’s residents, regardless of factors such as religion, caste, gender, or socio-economic status.

Aadhaar uses data provided by 10 finger scans and two iris scans and also applies stringent quality controls at the point of registration (Das, 2015). According to reports by UIDAI, combining the 12 measurements results in a low biometric failure-to enroll ratio of 0.14 percent, even in a population where many rural and manual workers are not able to provide high-quality fingerprints. The probability that a duplicate entry will not be caught (a false negative) was estimated at only 0.035 percent. The probability that an entry would be erroneously classified as a duplicate (a false positive), against the sample enrollment of 84 million in 2013 at the time of the study, was estimated at 0.057 percent. These numbers show the relative efficiency of Aadhaar from the state side of the implementation process.

Furthermore, according to several studies, Aadhaar has achieved the first stage of large, national level inclusion; since its rollout in 2008 the program has generated biometric IDs for about 95% of the Indian population (Das, 2015; Banerjee, 2016; Totapally, 2019). However, in a country as large and diverse as India, there are inevitably gaps in coverage. The 2019 *State of Aadhaar* report highlights regional and demographic gaps and mentions how Aadhaar might not have reached groups such as rural women, third gender individuals, and the homeless. The report further mentions how a lack of household-level data makes it difficult to effectively address these disparities (Totapally, 2019). The second stage in achieving national inclusion is addressing these gaps by assessing coverage among those who have been excluded, determining barriers in coverage, and ensuring every individual can capitalize upon the breadth of services provided by Aadhaar. My study contributes to filling this gap by using econometric analysis to help determine those who have been excluded and analyzing Aadhaar's effectiveness.

Although academic debates around Aadhaar are polarized, national household data has been thin on what Aadhaar has done for the residents of India. Since Aadhaar's rollout in 2009, few on-the-ground studies have been conducted to gather primary data and examine the program's impact on a diversity of subgroups. Literature and academic perspectives are also heavily polarized along lines of whether or not the program truly benefits India's marginalized groups, and what the Indian government's true intentions are behind collecting a massive dataset of information on nearly every Indian resident.

The *State of Aadhaar 2019* report is the largest primary dataset on the use of Aadhaar. The dataset I use for analysis in this paper comes from the *State of Aadhaar 2019* report, and the findings of this report are central to the intuition behind my econometric model. I therefore summarize some of the key findings in this section.

The *State of Aadhaar 2019* report, issued by the Dalberg Advisory Group and Omidyar Network, captures the perspectives of over 167,000 Indian residents through two national household surveys on Aadhaar and human centered design research. The report can be broken down into four sections: 1) the process of enrolling and updating one's Aadhaar ID, 2) using Aadhaar, 3) perceptions, satisfaction, and trust, and 4) the Aadhaar experience for different population groups. In terms of services and welfare provisions linked to Aadhaar, the report emphasizes the breadth and depth of Aadhaaar's role in a multitude of aspects of daily life: food, livelihood, energy, finance, education, and communication. Figure 1 below shows the extent of Aadhaar's role in a variety of sectors. Given that Aadhaar is becoming increasingly prominent in virtually every sector, it is important to examine Aadhaar's effectiveness in each of these sectors.

**Figure 1**

| | |
|---|---|
| Food | • 39% of people<br>(an estimated 330 million) give their Aadhaar-linked biometrics regularly to receive rations<br>• 29% of households (an estimated 80 million) depend on these rations for more than half of their monthly supplies |
| Livelihood | • 9 million people older than 65 years verified their Aadhaar to continue drawing social pensions, which for most is their only source of income<br>• 90% of farmers (an estimated 96 million) have given their Aadhaar at least once for their fertilizer subsidy<br>• 102 million MGNREGA workers have given their Aadhaar at least once in order to receive MGNREGA wages |
| Energy | • 50% of households have given their Aadhaar at least once for LPG subsidy<br>• 11% of households (an estimated 32 million) have given their Aadhaar at least once for kerosene subsidy |
| Finance | • 72% of adults (an estimated 609 million) have linked their bank accounts to Aadhaar; 87% of all bank accounts are linked<br>• 29% of household transactions are completed via the Aadhaar- enabled Payment System (AePS) |
| Education | • 47% of children (an estimated 125 million) enrolled in school using their own or their parent's Aadhaar |
| Communication | • 41% of adults (an estimated 345 million) gave their Aadhaar for a SIM connection |

Despite the high percentages shown in Figure 1, there are significant gaps in Aadhaar's coverage. The report states that 8% of India's population, approximately 103 million people, are not yet covered by Aadhaar (Totapally, 2019). Geographically, the report observes that the Indian states of Assam and Meghalaya are significantly behind on enrollment and take-up of Aadhaar, and rural regions fall significantly behind urban as well. A majority of these individuals who were not covered by Aadhaar stated that they wished to enroll but were not able to.

Furthermore, when compared to the national averages in each category, minority groups experienced a more prevalent lack of Aadhaar within the categories of age, gender, religion, and caste. The minorities within these groups were individuals aged 0-5, third gender individuals, Christians and Muslims, and scheduled tribes and scheduled castes, respectively. The report also finds that these marginalized groups are more likely to encounter errors on their Aadhaar cards, and that errors can result in exclusion from services. Therefore, demographic characteristics are an important metric of analysis when determining which populations are excluded from Aadhaar and the services Aadhaar provides. This understanding is central to my study and the construction of my econometric model.

### 3.3 Benefits and Pitfalls: Aadhaar's Implementation in India

Academic studies of Aadhaar's practical implementation and impact support the conclusions drawn by the *State of Aadhaar 2019* report and by Gelb and Clark (2013). Studies across India show that differing experiences with Aadhaar generally vary along demographic lines. For instance, Bhatti (2012) examines Aadhaar's linkage to the National Rural Employment Guarantee Act (MGNREGA), an Indian welfare program that provides employment and facilitates wage payments for rural Indian workers. Bhatti (2012) employs pilot studies administered in the Indian state of Jharkhand for his study. The study finds that illiteracy and lack of prior experience with financial products and services does not allow rural workers to take full advantage of Aadhaar's offerings. Bhatti proposes a certain model of financial service provision called the business-correspondent model, which will allow an in-person individual to assist rural workers with transactions. This study reveals that literacy, educational status, and urban vs. rural location are all significant factors that shape one's positive or negative experience with Aadhaar.

Drèze (2017) conducted a study examining the Public Distribution System (PDS) in India, a ration and resource distribution system, by using preexisting data from the Indian state of Jharkhand. The study examines issues related to Aadhaar and PDS linkage, including exclusion problems, transaction costs, and its impact on corruption. The study concludes that there are serious exclusion problems, particularly for vulnerable groups such as widows, the elderly and manual workers, as well as higher transaction costs. This study reinforces the findings of previous research by showing that marginalized and vulnerable groups face the most difficulty accessing Aadhaar and taking advantage of its services.

Biswas (2017) also conducted a study in Jharkhand that studies the impact of Aadhaar's linkage to the pension distribution system for old-age and widow pension beneficiaries. Biswas first highlights the pre-existing shortcomings in the pension distribution system before claiming that Aadhaar has only created additional problems for beneficiaries. Many elderly individuals are frail, illiterate, and not technologically savvy, resulting in them being unable to adapt to the Aadhaar-linked pension system. Many are also left out due to technical issues and inability to enroll, which means elders are unable to receive the money they need to survive. Biswas also notes that Aadhaar did not reduce any inefficiencies in the distribution system, thus leaving the reader with the impression that Aadhaar does not benefit the pension beneficiaries nor the government body that distributes the pensions.

Khera (2017) strongly concludes through her study that Aadhaar is not useful in the sense of being linked to Indian welfare programs unless major implementation issues are addressed. Khera's study relies on quantitative data from primary field studies, secondary data from government portals, figures obtained through queries made under the Right to Information Act, and responses to questions in Parliament to examine Aadhaar's impact on Indian welfare programs following the linkage of Aadhaar to three major programs. Khera states that although Aadhaar claims to decrease corruption and lower costs, data indicates that it does not do so. Additionally, marginalized groups are facing even higher levels of exclusion due to Aadhaar being made mandatory.

Lastly, several sociological studies analyze Aadhaar on a broader, political scale, thus contributing to the question of Aadhaar's role in policy and development. A study by Chaudhuri and König (2018) takes a sociological approach to analyze how Aadhaar shifts perceptions of citizenship in India. The paper states that Aadhaar maintains the neo-liberal notion of governing and concludes that Aadhaar transforms citizens into customers by facilitating the linkage to both public and private sector services. A study by Henne (2019) examines Aadhaar's role in fixing "leaking" systems in India and concludes that Aadhaar has further deteriorated an already crumbling system of governance by institutionalizing "leakiness". These two studies raise questions of whether connecting people to state and corporate systems via Aadhaar is an effective means of growth and development, especially given that many state and corporate systems in India battle corruption and major inefficiencies. These studies provide a framework for interpreting the implications that arise from my paper. Although my study examines whether

Aadhaar is successful in enabling individuals to access public and private services, it is important to keep the bigger picture in mind: do these public and private services enhance the lives of Indians? This question may be a topic for further research.

Although the research captured in this literature review reveals relevant findings about Aadhaar and its implementation, these studies do not conduct econometric analysis to unearth mathematical relationships between Aadhaar and certain desired outcomes. In essence, most of the studies captured here in Section 3 serve as case studies of Aadhaar's practical implementation and lived experiences with the program. My study therefore builds upon the findings of this prior research to construct and execute an econometric model. My model will describe the mathematical relationship between enrollment in Aadhaar and access to four key public and private services in India: 1) the Public Distribution System (PDS), 2) the National Rural Employment Guarantee Act (MGNREGA), 3) bank accounts (savings accounts), and 4) SIM cards. The next subsections of the literature review provide brief background about each of these selected public and private services. This background and intuition are then used to construct my econometric models, as detailed in Section 5.

### 3.4. The Public Distribution System (PDS)

The Public Distribution System (PDS), first established in 1945, is India's most far-reaching, in terms of coverage, and the most expensive public program (Ahluwalia, 1993). The program plays a crucial role in reducing food insecurity in India. PDS works alongside a free market and strives to ensure food security for millions of consumers across the country by subsidizing rations of rice, wheat, sugar and edible oil through thousands of retail outlets spread across the country (Ahluwalia, 1993). First, the essential commodities are procured by the central government from private producers at below market prices, and secondly, they are redistributed through ration shops throughout the nation (Masiero, 2020).

The PDS was reformed in 1997, as India endured a series of neoliberal reforms and structural adjustments. As the PDS came under the scrutiny of World Bank economists, inefficiencies and leakages were pinpointed and a reduction in public expenditures was suggested. The PDS system was therefore reformed to be more "targeted." The new system created new standards for determining society's most needy members; households that did not qualify opted out of the PDS system and returned to buying essential commodities off of the free market.

In terms of challenges, systematic leakage constitutes a serious issue in the program and is the principal disruption to its normal functioning (Masiero, 2020). The integration of Aadhaar and PDS was intended to confront these leakages by deduplicating IDs, thus ensuring that government expenditures and rations were effectively being directed to those who need it most. Incorporation of Aadhaar in the PDS started in 2011-2012, with pilot projects undertaken at the local level in different states. These projects introduced Aadhaar recognition in the ration shop space that constitutes the last mile of the supply chain. Although a Supreme Court order in 2013 prohibited making Aadhaar mandatory for accessing anti-poverty programs, many beneficiaries and frontline workers distributing PDS rations falsely believe Aadhaar is mandatory for distribution. This constitutes a large problem in the current on-the-ground-implementation of PDS via Aadhaar.

According to a study by George and McKay (2019), the PDS system subsidizes and provides food grains to over 800 million people, or approximately two thirds of India's population (George and McKay, 2019). The *State of Aadhaar 2019* report also stated that 39% of

Aadhaar-holders have their ID linked to PDS, and 29% of households depend on these rations for more than half of their monthly supplies (Totapally, 2019). The PDS was therefore selected as a key public service for this paper given it scale and impact in India.

The next subsections examine the four services selected for analysis in this paper and describe the scale and impact of each service. Shortcomings of each service are also highlighted.

### 3.4.b The National Rural Employment Guarantee Act (MGNREGA)

The MGNREGA is arguably the largest public workfare program in the world. It has generated more than 22.68 billion person-days of work, involving expenditures of Indian rupees 3776.7 billion (USD 58 billion) since its inception in 2006. The Act was passed on September 5, 2005, with a stated goal to improve livelihood security for rural households by providing up to 100 days of guaranteed wage employment in every financial year to each household whose adult members volunteer to do unskilled manual work (Government of India, 2013). Some examples of permissible work, typically provided within the village, include water conservation and harvesting, land development, horticulture and plantations, and rural connectivity. Workers are paid according to a schedule of rates established by state governments for different tasks performed in different soil conditions. The program had a phased-in rollout starting with the 200 districts deemed the poorest, and then gradually expanded to cover all of India's districts between 2006 and 2008.

A study by Narayanan et. al (2017) examined the decline of MGNREGA starting in 2009, when the program declined in terms of both total expenditure and person-days of employment. Narayan et. al outline how some scholars suggest that the programs needs were met, and that a more targeted program induced many people to select out. Others argue that there is an unmet demand for work, and that workers are actively feeling discouraged to seek out work.

The *State of Aadhaar 2019* report states that 102 million MGNREGA workers have given their Aadhaar at least once to receive their wages. The importance of the MGNREGA program and the statistics on its integration with Aadhaar reveal the scale and impact of the program, thus deeming it one of the most important anti-poverty public programs in India.

### 3.4.c SIM cards and Bank Accounts

The two private services selected are bank account access and SIM card access. In this paper, access to bank accounts is measured through a variable indicating usage of a savings account. By measuring access to savings accounts, I am able to better gage whether individuals are actually utilizing their bank accounts.

A multitude of pre-existing literature cites the importance of bank accounts for building savings and growing a family's economic, social, and political opportunities.[1] Financial inclusion is the broad-based delivery of banking and other financial services at affordable cost to the poorest sections of society. In India, financial inclusion emphasizes the inclusion of the maximum number of people under formal financial systems. The most important part of financial services in a region is typically measured by the number of people who have access to bank accounts.

Financial inclusion and access to bank accounts allows individuals to securely accumulate, hold, and transfer capital. This provides the ability to cope with shocks and ensures a safety net for future expenditures. A multitude of studies cite how financial security improves the general wellbeing of individuals, households, and communities at large, and contributes to economic development on a national scale (Gupte et al. (2012), Leeladhar, V. (2006), and Sarma, M., & Pais, J. (2011)).

The Indian government has undertaken ambitious strategies for including financial inclusion in its national agenda. With the increase in technology-driven branchless banking initiatives, government strategies have been largely successful in regard to extending access; nearly 60% of the Indian population is banked (Morawczynski et al, 2010). However, empirical evidence suggests that the majority of bank accounts are not being utilized, especially not by the poor who are the target of financial inclusion (Morawczynski et al, 2010).

A study by Morawczynski et al, 2010 examines the reasons behind low bank account usage, and pinpoints several population groups that face the most difficulty in taking advantage of financial services. In summary, the three key demographic groups that face difficulty accessing financial services are low-wage earners, low-earning and MGNREGA-dependent workers, and individuals in remote areas that do not have financial inclusion related infrastructure. This demographic information is key to the construction of my econometric model evaluating access to bank accounts.

Along the same lines of financial inclusion, access to SIM cards is increasingly important in a globalized world. SIM card access affects livelihoods, health, education, and more. In developing economies such as India, especially in rural parts of India, mobile connectivity helps drive economic growth, fostering business development and wider market access. Mobile phones provide a more reliable alternative to roadways and postal systems for communities in remote and underserved areas. As stated by Anand et al. (2012), “In developing countries, [mobile phones] are creating opportunities for users to access market information, monitor health care, transfer money and promote literacy.”

Additionally, previous studies cite the importance of SIM cards and cellular phones in today’s increasingly technology-oriented world (Kendall and Voorhies (2014) and Klonner and

Nolen (2010)). For instance, Anand et al. (2012) state that a lack of transportation, high illiteracy levels, and migrant labor are some of the characteristics of rural areas that emphasize the need for real-time voice communication in rural areas. Additionally, their argument is founded in the belief that that voice communication in the developing world is an enabler of political freedom, economic growth, and efficient health care. Therefore, given the established importance of savings accounts and SIM cards, these two variables were selected as dependent variables for analysis.

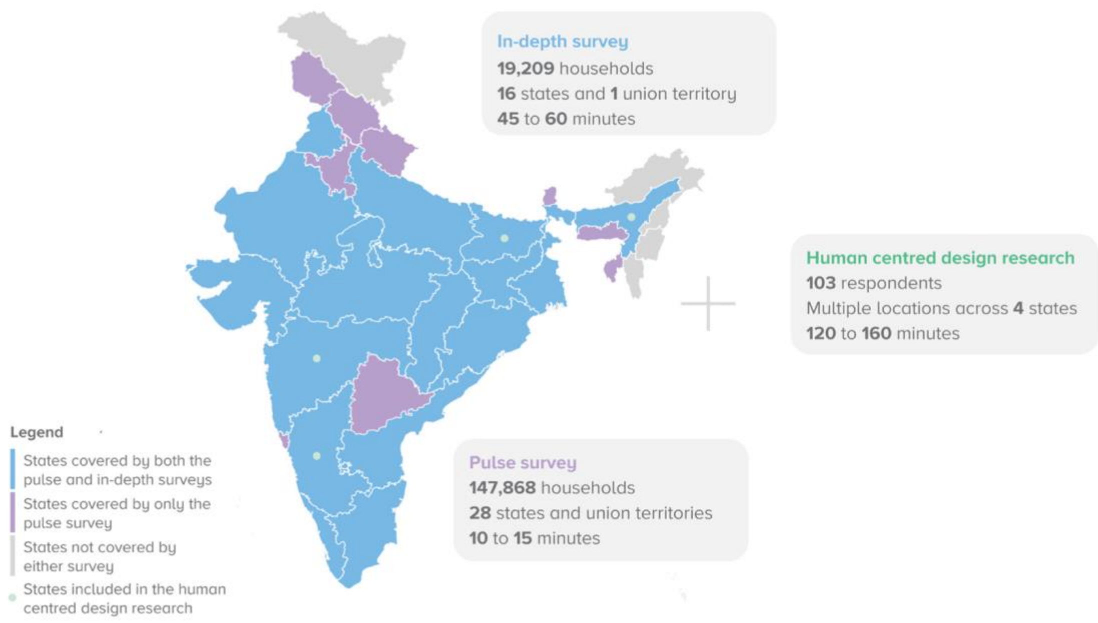
Literature on the selected services reveals how these four public and private services are important in terms of scope and impact. Public programs such as PDS and MGNREGA impact the lives and livelihoods of millions of people across India, while savings accounts and SIM cards are now important precursors to social and economic mobility. The analysis in this paper relies on the findings of preexisting literature to examine how marginalized and vulnerable groups are able to access these key selected services.

### 4.1 Data Source

Data for this study comes from the *State of Aadhaar 2019* report. As stated previously in Section 3, the *State of Aadhaar 2019* report captures the perspectives of over 167,000 Indian residents through two national household surveys on Aadhaar and human centered design research. This report is the largest primary dataset on Aadhaar. These two national surveys include: 1) a pulse survey of a panel of 147,868 households in 28 states and union territories, and 2) an in-depth 45-minute survey of 19,209 households across 16 states and 1 union territory. Coverage of the research, as published by the Dalberg Advisory Group and Omidyar Network, is shown below in Figure 2.

The Dalberg Advisory Group and Omidyar Network provide six datasets that cover household and individual level surveys. For this project I use two datasets in order to capture demographic information as well as information about usage of the four public and private services.

**Figure 2**




The demographic information comes from an individual-level survey of 17,112 individuals and contains 37 variables. The information about usage of public and private services comes from a household-level survey of 4,793 households and includes a detailed questionnaire entailing 1,852 variables. All of these variables, in both the demographics and indicators datasets, are categorical or binary variables.

For my analysis, I chose to include a variable measuring access to a savings account rather than access to a bank account. A variable measuring savings account usage implies whether or not an individual is actually making use of their bank account. This addresses the issue raised in a study by Morawczynski et al (2010), where low-income individuals possess bank accounts but are not using them.

**4.2 Data Preparation**

In order to merge the two datasets, I compile the demographics dataset with the indicators dataset by using the unique identification number; this number is the same across the household level and individual level datasets, allowing me to merge the data accordingly. However, in the individual-level dataset, every individual who was a part of the same household had the same identification number. Therefore, in order to merge the data, I relied on the “member number” variable and only used the first member of every household. My final dataset used for analysis essentially entails demographic information and indicator information for the head of every household.

While this method of data cleaning allows me to match and merge data effectively, my analysis is relying solely upon information about the head of each household. Typically, the head of household in an average Indian home is an older male. Before merging the datasets, I completed a baseline check of the heads of households to ensure there was enough variation in the data. However, the fact that the final dataset is solely covering the head of each household is a potential shortcoming of this analysis and should be kept in mind.

Majority of the variables chosen for analysis in this paper are categorical variables. The variables measuring usage of the four selected public and private services include three categories: “Yes”, “No”, and “Don’t Know.” There were also a small number of missing observations for each of these variables. The missing observations and “Don’t Know” observations were not correlated with any of the independent variables, so I omitted these two categories from each of the indicator variables. The Aadhaar variable also contained a “Don’t

Know" category and missing variables. These two categories were not correlated with any independent variables and were therefore omitted.

I also omitted categories from several demographic variables in cases where there were few observations for that category and when that category was not correlated with any independent variables. Omitting categories for demographic and indicator variables entailed removing those observations entirely from the final dataset used for analysis.

Finally, I releveled the variables in order to make the data most efficient for analysis. For instance, I re-leveled the Aadhaar variable so that the "No" category is omitted in the regression.

### 4.3 Summary Statistics

The final, merged dataset used for analysis contains 4,511 observations and 12 variables. For the purpose of this paper, I selected 7 demographic variables from the larger *State of Aadhaar* dataset for analysis. This data is captured at the individual level and includes: 1) a unique identification number for each observation, 2) gender, 3) caste, 4) state of residence, 5) urban vs. rural, 6) migrant laborer, and 7) a member ID number that indicates which household member number they are.  I also selected 5 variables, captured at the household level, that capture usage of Aadhaar and the selected services. These variables are: 1) a unique identification number for each observation, 2) Aadhaar, 3) PDS card, 3) SIM card, 4) MGNREGA card, and 5) savings account.

In terms of the demographic information, the "caste" variable captures information about caste denominations of each observation through six categories: 1) General, 2) No response, 3) other backward class (OBC), 4) non-vulnerable scheduled tribe (ST), 5) scheduled caste (SC), and 6) vulnerable ST. "General" denotes castes whose members are on average ahead of other Indians socially and economically. Scheduled tribes (ST) and scheduled castes (SC) are the most disadvantaged caste denominations; SCs face social and economic isolation and STs are marginalized on the basis of geographic location. Other backward classes (OBC) are groups that are historically marginalized in India and continue to face oppression but are not ST or SC.

The *Gender* variable includes three categories (0 = male, 1 = female, 2 = third gender). The following variables are binary: *Migrant Laborer* (0 = not a migrant worker, 1 = migrant worker), *UrbanRural* (0 = urban, 1 = rural), Aadhaar (0 = no Aadhaar, 1 = Aadhaar). The indicators, *PDS*, *MGNREGA*, *Savings Accounts*, and *SIM cards*, are all binary (0 = no usage, 1 = usage).

Out of the 4,511 observations, 89% have access to Aadhaar and 11% do not. In terms of access to the selected services, 61% use PDS and 39% do not; 14% use MGNREGA and 86% do not; 62% use a savings account and 38% do not; 58% use SIM cards and 42% do not. These statistics are generally representative of Indian society at large and usage of these selected services.

In terms of *Gender*, 68% of the sample is male; 23% is female; 9% is third gender. In terms of *UrbanRural*, 61% of the sample is located in urban areas and 39% in rural areas. In terms of *Migrant Worker*, 94% of the sample are not migrant workers and 6% are migrant workers. For the *Caste* variable, 30% of the sample classifies as Other Backwards Classes (OBC) and 27% classifies as Scheduled Caste (SC); 19% classify as general; “No Response”, non-vulnerable Scheduled Tribe, and Scheduled Tribe comprise approximately 7-8% of the sample each. These numbers are not completely accurate representations of Indian society as a whole. For instance, the sample contains a significantly larger number of male observations than female. However, the sample data does provide good variation across demographic groups.

## 4. Empirical Approach

I use two empirical approaches to analyze the effect of Aadhaar attainment on ability to access four key public and private services in India. Each approach involves four separate regression equations, one for each of the indicators: 1) PDS, 2) MGNREGA, 3) SIM cards, and 4) Savings accounts.

For my first approach, I regress the variable capturing usage of the indicator (1= access, 0= no access) on Aadhaar attainment (1= enrolled, 0= not enrolled) and a set of demographic controls.The inclusion of demographic controls is informed by the findings of prior research on each of the indicators. These controls are intended to shed light on which external demographic factors affect the probability of using the key public and private services.

I have four separate equations - one for each indicator. I run a logit and probit analysis on each of these equations, and analyze the coefficients using average marginal effects and by interpreting the regression coefficients as probabilities. Coefficients are interpreted as probabilities through the formula:

$$\frac{exp(\beta)}{1 + exp(\beta)}$$

Results from the probit analysis and a comparison between the logit and probit outputs of this strategy are included in the Appendix. The four equations used in the first approach are as follows:

**Approach 1**

**Equation 1: PDS**

$$PDS_i = \beta_0 + \beta_1 Aadhaar_i + \beta_2 Gender_i + \beta_3 UrbanRural_i + \beta_4 Caste_i + u$$

**Equation 2: MGNREGA**

$$MGNREGA_i = \beta_0 + \beta_1 Aadhaar_i + \beta_2 Gender_i + \beta_3 UrbanRural_i + \beta_4 Caste_i + \beta_5 MigrantLaborer_i + u$$

**Equation 3: Savings account**

$$SavingsAccount_i = \beta_0 + \beta_1 Aadhaar_i + \beta_2 Gender_i + \beta_3 UrbanRural + \beta_4 Caste_i + \beta_5 MigrantLaborer_i + u$$

**Equation 4: SIM card**

$$SIMcard_i = \beta_0 + \beta_1 Aadhaar_i + \beta_2 Gender_i + \beta_3 UrbanRural_i + \beta_4 Caste_i + u$$

---

By analyzing a coefficient using marginal effects, I am able to gage the rate at which access to an indicator changes when the relevant independent variable changes, with the rest of the covariates held constant. Average marginal effects calculate marginal effects (AMEs) at every observed value of X, and average across the resulting effect estimates. AMEs are a particularly useful tool of analysis because they produce a single quantity that reflects the full distribution of X and better capture variability.

For this first approach, I hypothesize that the coefficients on marginalized or vulnerable groups will be negative. For instance, for the *Gender* variable, I hypothesize that the coefficient on Third Gender and Female will be negative, since women and third gender individuals are more marginalized demographic groups in India as compared to males. In a similar vein, I hypothesize that the coefficients on tbe *Scheduled Tribe* and *Scheduled Caste* categories will be negative since these groups are more vulnerable as compared to individuals in the *General* category. Vulnerable or marginalized groups are less likely to be able to access public or private services such as PDS, MGNREGA, SIM cards, and savings accounts; therefore, I hypothesize negative coefficients.

For my second approach, I utilize the same equations as above but include interaction terms in each of the equations. In each equation I added two interaction terms, one that analyzes the interaction between *Aadhaar* and *Gender*, and the other analyzes the interaction between *Aadhaar* and *UrbanRural*. Similar to the first strategy, I run a logit and probit analysis on each of these equations. Results from the probit analysis and a comparison between the logit and probit outputs of this strategy is included in the Appendix.

The four equations for this second approach are as follows:

**Approach 2**

**Equation 1: PDS**

$$PDS_i = \beta_0 + \beta_1 Aadhaar_i + \beta_2 Gender_i + \beta_3 UrbanRural_i + \beta_4 Caste_i + \beta_5 (Aadhaar*Gender)_i + \beta_6 (Aadhaar*UrbanRural)_i + u$$

**Equation 2: MGNREGA**

$$MGNREGA_i = \beta_0 + \beta_1 Aadhaar_i + \beta_2 Gender_i + \beta_3 UrbanRural_i + \beta_4 Caste_i + \beta_5 MigrantLaborer_i + \beta_6 (Aadhaar*Gender)_i + \beta_7 (Aadhaar*UrbanRural)_i + u$$

**Equation 3: Savings account**

$$SavingsAccount_i = \beta_0 + \beta_1 Aadhaar_i + \beta_2 Gender_i + \beta_3 UrbanRural_i + \beta_4 Caste_i + \beta_5 MigrantLaborer_i + \beta_6 (Aadhaar*Gender)_i + \beta_7 (Aadhaar*UrbanRural)_i + u$$

**Equation 4: SIM card**

$$SIMcard_i = \beta_0 + \beta_1 Aadhaar_i + \beta_2 Gender_i + \beta_3 UrbanRural_i + \beta_4 Caste_i + \beta_5 (Aadhaar*Gender)_i + \beta_6 (Aadhaar*UrbanRural)_i + u$$

The inclusion of interaction terms tests the hypothesis that the relationship between certain demographic characteristics and access to public and private services is different for individuals with Aadhaar and without Aadhaar. The interaction term between *Aadhaar* and *Gender* tests whether the relationship between access to public and private services and being a woman or third gender individual is different for individuals with Aadhaar and without Aadhaar. The interaction term between *Aadhaar* and *UrbanRural* tests whether the relationship between access to public and private services and being located in rural India or urban India is different for individuals with Aadhaar and without Aadhaar.

I selected *Gender* and *UrbanRural* as the two demographic variables to interact with Aadhaar based on current findings on the shortcomings of Aadhaar. Several studies, such as Khera (2017) and Totapally (2019), find that individuals in rural India face higher levels of difficulty in accessing Aadhaar and taking advantage of the services it provides. This is due to the difficulty that arises accessing remote rural areas, as well as demographic trends such as low literacy and lower income levels that are characteristic of rural India (Khera, 2017). According to the objective of the Aadhaar program, access to Aadhaar should improve the ability of individuals in rural India to access public services such as PDS and MGNREGA and private services such as SIM cards and savings accounts.

Totapally (2019) finds that third gender individuals face very high levels of difficulty in accessing Aadhaar and taking advantage of the services it provides. This is due largely to the marginalization and discrimination that third gender individuals already face in Indian society. Aadhaar, as a program, aims to enable individuals of all backgrounds and identities to access public and private services. I therefore hypothesize that the coefficients on both of these interaction terms will be positive and statistically significant in all of the four regression equations.

In both of the approaches, I hypothesize that the coefficient on *Aadhaar* will be positive. A positive coefficient indicates that individuals with Aadhaar are more likely to be able to access certain services as compared to individuals without Aadhaar. This hypothesis is in line with the objective of the Aadhaar program, as put forward by the Unique Identification Authority of India.

## 6. Results

### 6. 1 Approach 1

Table 1 the Appendix includes the results from Approach 1: regressing the variables capturing usage of the chosen services on Aadhaar attainment and a set of demographic controls. I find that the coefficients on Aadhaar are positive and statistically significant in all of the equations. This means that individuals with access to Aadhaar are more likely to be able to access public and private services than individuals without Aadhaar. Through converting the logit coefficients to probabilities, the probabilities of individuals with Aadhaar using the public or private service are conveyed. The probability of an individual with Aadhaar using PDS is 87%; using MGNREGA is 65%; using a savings account is 83%; using a SIM card is 76%. These numbers reveal that individuals with Aadhaar are generally likely to be connected to state and private services.

The average marginal effect (AME) on Aadhaar is positive across all of the equations, revealing that, all else being equal, the probability of having access to the indicator increases for individuals with Aadhaar. For instance, in terms of PDS, the AME on Aadhaar is .41. This means the probability of having access to PDS increases by 41 percentage points for individuals with Aadhaar. The AME on Aadhaar is lowest for MGNREGA; the probability of having access to MGNREGA increases by 5.83 percentage points for individuals with Aadhaar.

In terms of the demographic controls, the findings of Table 1 are generally in line with my hypothesis. The coefficients on *Third Gender* are statistically significant and negative for PDS and MGNREGA. Looking at the AMEs, this reveals the probability of having access to PDS decreases by 31 percentage points for third gender individuals, and the probability of having access to MGNREGA decreases by 13 percentage points. This is intuitive because third gender individuals are among the most marginalized groups in India, thus increasing the barriers that stand in the way of accessing the selected services.

Interestingly, the coefficients on *Female*, are positive for PDS and savings accounts. In terms of PDS, this could be due to the fact that PDS rations are received by households rather than individuals. Since women are traditionally responsible for household duties such as water collection, collecting rations could fall under the realm of female responsibilities. This theory could explain the positive coefficient on *Female*, which reveals that women are more likely than men to access PDS. In the case of savings accounts, the positive coefficient is contrary to my hypothesis. A positive coefficient indicates that women are more likely than men to access savings accounts. This finding contrasts previous findings, such as those detailed by Ghosh, S. and Vinod, D. (2017), that finds that women are far behind men in terms of accessing financial services. It is possible that the women captured by the sample data are better connected to financial service providers or just naturally more likely to access savings accounts.

Furthermore, the coefficients on *Rural* are all statistically significant and are positive for each service except for SIM cards. This is counterintuitive and does not support my hypothesis that rurally located individuals are less likely to have access to public and private services. In terms of the AME, the AME on *Rural* is .08 for both PDS and savings accounts. This means the probability of having access to PDS or a savings account increases by 8 percentage points for rurally located individuals. In terms of Caste, the vulnerable groups had negative coefficients, which supports my hypothesis that marginalized castes are less likely to have access to public and private services.

In summary, the findings of Table 1 reveal that individuals with access to Aadhaar are more likely to use PDS, MGNREGA, savings accounts, and SIM cards. The findings also show that marginalized and vulnerable demographic groups are less likely to be able to use the selected services as compared to their more privileged counterparts. Although there are some outliers, these findings are in line with my hypothesis.

Examining the log likelihood values from Tables 1 and 2 reveals that the probit and logit models for Approach 1 are both qualitatively similar. McFadden $R^2$ values were also calculated for each of the equations in the logit and probit models. Results of the McFadden $R^2$ calculations are discussed further in section 6.2 and are included in Table 10.

## 6.2 Approach 2

Table 3 in the Appendix includes the results from Approach 2: regressing the variables capturing usage of the indicators on Aadhaar attainment and including interaction terms between *Aadhaar* and *Gender* and *UrbanRural*. Looking at the interaction terms, the coefficients on the interaction between *Third Gender* and *Aadhaar* are statistically significant and negative for PDS, MGNREGA, and SIM cards. This means third gender individuals with Aadhaar have a negative probability of having access to PDS, MGNREGA, and SIM cards as compared to non-third gender individuals with Aadhaar. This finding contrasts my hypothesis and the objective of the

Aadhaar program. Access to Aadhaar is intended to help overcome barriers that marginalized individuals, such as third gender individuals, might otherwise face.

The significant coefficients on the interaction between *Aadhaar* and *Rural* are also negative. This reveals that rurally located individuals with Aadhaar have a negative probability of having access to MGNREGA, savings accounts, and SIM cards as compared to rurally located individuals without Aadhaar. Again, this contrasts my hypothesis and the objective of the Aadhaar program.

Khera (2017) and other academics point out potential explanations for these findings in their research. Due to misinformation about Aadhaar, individuals are sometimes unable to access certain public services. For instance, although having or showing your Aadhaar card is not mandatory for receiving one's PDS rations, people might falsely believe that they must have Aadhaar to access services. Aadhaar also has high error rates among vulnerable and marginalized groups, such as third gender individuals or the rural poor.

Additionally, the coefficients on *Female*, *Third Gender*, and *Rural* change once the interaction terms are included. For instance, the coefficient on *Female* in equation (3) evaluating savings accounts, although it is no longer statistically significant, is now negative rather than positive. With the inclusion of an interaction term, the coefficient on *Female* is interpreted as the effect of being a female when Aadhaar = 0. The negative coefficient means females are less likely than males to use savings accounts. In contrast, the coefficient on *Rural* in equation (4) evaluating SIM cards is now positive in Table 3 as compared to Table 1. Here, the coefficient on *Rural* is interpreted as the effect of being rurally located on using a SIM card when Aadhaar = 0. A positive coefficient means individuals in rural locations are more likely to use SIM cards than individuals in urban locations. These changes in coefficients are the result of interaction terms, which indicate that the effect of *Gender* and *UrbanRural* on the selected public and private services are different for different values of *Aadhaar*.

Examining the log likelihood values from Tables 3 and 4 reveals that the probit and logit models for Approach 2 are both qualitatively similar. The McFadden $R^2$ was also calculated for each equation using the following formula:

$$1 - \frac{log\ likelihood(model)}{log\ likelihood(model\ with\ no\ covariates)}$$

The McFadden $R^2$ of each equation is reported in Table 10 of the Appendix. The closer the McFadden $R^2$ value is to 1, the more well-fitted model the model is. The results of Table 10, show that the model estimated access to MGNREGA is the best-fitted model, since these values are the closest to 1 at approximately .15. The model estimating access to SIM cards is the least well-fitted model, since these values are the closest to 0 at .04.

**Conclusion**

While this analysis shows a positive relationship between Aadhaar attainment and access to public and private services in India, it is important to take reverse causality into account. Although the research question and model employed in this paper argue that Aadhaar enables individuals to gain access to certain public and private services, it is possible that access to the public and private services is, in reality, prompting access to Aadhaar. This could be plausible since infrastructure for public services has been on the ground for much longer than Aadhaar. An individual who already uses PDS, for example, could have better access to the Aadhaar system once it was introduced and could be more familiar with state administrative processes.

The findings of this paper suggest that Aadhaar has a positive relationship with usage of public and private services in India. In Approach 1, the negative coefficients on marginalized groups such as *Third Gender* and *Vulnerable Scheduled Tribe* indicate that vulnerable groups are less likely to use the select services than their more privileged counterparts. In Approach 2, the majority negative coefficients on the interaction terms indicate that Aadhaar may do little to improve marginalized individuals' access to the select public and private services.

These findings are largely corroborated by the literature outlined in Section 3. As prior studies have pointed out, individuals at the bottom of the ladder are not only less likely to be able to obtain Aadhaar IDs, but they are often not empowered with knowledge on how to use this technological tool. As usage of biometric ID systems like Aadhaar continues to expand around the world, it is important to examine whether these systems are truly impacting every corner of society equally and equitably. This study intends to raise questions about how Aadhaar's implementation and linkage to public and private services in India can be improved in order to truly ensure social inclusivity for all of India.

**References**


Ahluwalia, D. (1993). Public distribution of food in India: Coverage, targeting and leakages. Food Policy, 18(1), 33-54.

Anand, A., Pejovic, V., Belding, E. M., & Johnson, D. L. (2012, March). VillageCell: Cost effective cellular connectivity in rural areas. In Proceedings of the Fifth International Conference on Information and Communication Technologies and Development (pp. 180-189).

Banerjee, S. (2016). Aadhaar: Digital inclusion and public services in India. World Development Report, 81-92.

Bhatti, B. (2012). Aadhaar-enabled payments for NREGA workers. Economic and Political Weekly, 16-19.

Biswas, S. (2017). Struggles of Pensioners in Jharkhand. *Economic & Political Weekly*, *52*(52), 30.

Chaudhuri, B., & König, L. (2018). The Aadhaar scheme: a cornerstone of a new citizenship regime in India?. Contemporary South Asia, 26(2), 127-142.

Das, S. K. (2015). Making the Poor Free?: India's Unique Identification Number. Oxford University Press.

Drèze, J et al. (2017). Aadhaar and food security in Jharkhand. Economic & Political Weekly, 52(50), 51.

Gelb, A., & Clark, J. (2013). Identification for development: the biometrics revolution. Center for Global Development Working Paper, (315).

George, N. A., & McKay, F. H. (2019). The Public Distribution System and Food Security in India. International journal of environmental research and public health, 16(17), 3221.

Ghosh, S. and Vinod, D. (2017). What constrains financial inclusion for women? Evidence from Indian micro data. World Development, 92, 60-81.

Gupte, R., Venkataramani, B., & Gupta, D. (2012). Computation of financial inclusion index for India. Procedia-Social and Behavioral Sciences, 37, 133-149.

Henne, K. (2019). Surveillance in the Name of Governance: Aadhaar as a Fix for Leaking Systems in India. In Information, Technology and Control in a Changing World (pp. 223-245). Palgrave Macmillan, Cham.

Kendall, J., & Voorhies, R. (2014). The Mobile-Finance Revolution; How Cell Phones Can Spur Development. Foreign Aff., 93, 9.

Khera, R. (2017). Impact of Aadhaar in welfare programmes. Available at SSRN 3045235. Sharma, Vikas. "Aadhaar-a unique identification number: Opportunities and challenges ahead."

Klonner, S., & Nolen, P. J. (2010). Cell phones and rural labor markets: Evidence from South Africa.

Leeladhar, V. (2006). Taking banking services to the common man-financial inclusion. Reserve Bank of India Bulletin, 60(1), 73-77.

Masiero, S. (2020). Biometric infrastructures and the Indian public distribution system. South Asia Multidisciplinary Academic Journal, (23).

Morawczynski, O., Hutchful, D., Cutrell, E., & Rangaswamy, N. (2010, December). The bank account is not enough: Examining strategies for financial inclusion in India. In Proceedings of the 4th ACM/IEEE international conference on information and communication technologies and development (pp. 1-11).

Narayanan, S., Das, U., Liu, Y., & Barrett, C. B. (2017). The “discouraged worker effect” in public works programs: Evidence from the MGNREGA in India. World Development, 100, 31-44.

Sarma, M., & Pais, J. (2011). Financial inclusion and development. Journal of international development, 23(5), 613-628.

Singh, R., & Jackson, S. J. (2017, May). From Margins to Seams: Imbrication, Inclusion, and Torque in the Aadhaar Identification Project. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 4776-4824).

Szreter, S. (2007). The right of registration: development, identity registration, and social security—a historical perspective. World Development, 35(1), 67-86.

Totapally, Sweta et al. (2019). *State of Aadhaar Report 2019*. Dalberg, 2019.

Unicef. (2005). Child poverty in rich countries 2005. UNICEF Innocenti Research Centre.

## Appendix

Table 1: Approach 1 - Logit Regression (log odds ratio)

| | Dependent variable: | | | |
|---|---|---|---|---|
| | PDS | MGREGA | Savings Account | SIM card |
| | (1) | (2) | (3) | (4) |
| Has Aadhaar | 1.886*** | 0.635*** | 1.578*** | 1.176*** |
| | (0.123) | (0.191) | (0.112) | (0.108) |
| Female | 0.474*** | -0.102 | 0.235*** | -0.282*** |
| | (0.082) | (0.107) | (0.079) | (0.074) |
| Third Gender | -1.386*** | -2.452*** | -0.179 | 0.444*** |
| | (0.131) | (0.589) | (0.119) | (0.123) |
| Rural | 0.415*** | 1.804*** | 0.384*** | -0.196*** |
| | (0.070) | (0.102) | (0.069) | (0.065) |
| Caste: No Response | -0.304** | 0.373 | -0.467*** | -0.548*** |
| | (0.147) | (0.227) | (0.139) | (0.135) |
| Caste: Non-vulnerable ST | -0.209 | 0.536*** | -0.389*** | -0.609*** |
| | (0.137) | (0.177) | (0.132) | (0.127) |
| Caste: OBC | -0.123 | -0.021 | -0.143 | 0.109 |
| | (0.098) | (0.145) | (0.096) | (0.093) |
| Caste: SC | -0.254** | 0.150 | -0.182* | -0.188** |
| | (0.100) | (0.144) | (0.098) | (0.094) |
| Caste: Vulnerable ST | -0.052 | 0.786*** | -0.591*** | -0.480*** |
| | (0.148) | (0.180) | (0.140) | (0.136) |
| Migrant Laborer | | 0.209 | -0.175 | |
| | | (0.214) | (0.138) | |
| Constant | -1.209*** | -3.488*** | -0.861*** | -0.470*** |
| | (0.141) | (0.227) | (0.131) | (0.126) |
| Observations | 4,511 | 4,511 | 4,511 | 4,511 |
| Log Likelihood | -2,686.814 | -1,568.431 | -2,809.312 | -2,944.082 |
| Akaike Inf. Crit. | 5,393.628 | 3,158.863 | 5,640.625 | 5,908.165 |

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2: Approach 1 - Probit Regression (log odds ratio)

| | Dependent variable: | | | |
|---|---|---|---|---|
| | PDS (1) | MGREGA (2) | Savings Account (3) | SIM card (4) |
| Has Aadhaar | 1.135*** | 0.350*** | 0.980*** | 0.725*** |
| | (0.071) | (0.101) | (0.067) | (0.066) |
| Female | 0.283*** | -0.049 | 0.143*** | -0.175*** |
| | (0.049) | (0.059) | (0.048) | (0.046) |
| Third Gender | -0.831*** | -1.033*** | -0.110 | 0.269*** |
| | (0.078) | (0.220) | (0.073) | (0.074) |
| Rural | 0.249*** | 0.972*** | 0.233*** | -0.119*** |
| | (0.042) | (0.053) | (0.042) | (0.040) |
| Caste: No Response | -0.193** | 0.176 | -0.286*** | -0.338*** |
| | (0.088) | (0.123) | (0.085) | (0.083) |
| Caste: Non-vulnerable ST | -0.119 | 0.289*** | -0.234*** | -0.377*** |
| | (0.082) | (0.099) | (0.080) | (0.079) |
| Caste: OBC | -0.073 | -0.011 | -0.084 | 0.067 |
| | (0.059) | (0.078) | (0.058) | (0.057) |
| Caste: SC | -0.151** | 0.073 | -0.109* | -0.116** |
| | (0.060) | (0.078) | (0.059) | (0.058) |
| Caste: Vulnerable ST | -0.026 | 0.430*** | -0.358*** | -0.298*** |
| | (0.088) | (0.102) | (0.086) | (0.084) |
| Migrant Laborer | | 0.141 | -0.110 | |
| | | (0.112) | (0.085) | |
| Constant | -0.718*** | -1.958*** | -0.539*** | -0.287*** |
| | (0.083) | (0.118) | (0.079) | (0.077) |
| Observations | 4,511 | 4,511 | 4,511 | 4,511 |
| Log Likelihood | -2,687.943 | -1,569.462 | -2,808.997 | -2,944.182 |
| Akaike Inf. Crit. | 5,395.885 | 3,160.925 | 5,639.994 | 5,908.363 |

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3: Approach 2 - Logit Regression (log odds ratio)

| | Dependent variable: | | | |
|---|---|---|---|---|
| | PDS | MGREGA | Savings Account | SIM card |
| | (1) | (2) | (3) | (4) |
| Has Aadhaar | 2.098*** | 1.455*** | 2.190*** | 1.735*** |
| | (0.192) | (0.527) | (0.197) | (0.175) |
| Female | 0.982*** | -1.124** | -0.034 | -0.397 |
| | (0.260) | (0.488) | (0.276) | (0.276) |
| Third Gender | -0.694* | -1.183 | -0.096 | 0.996*** |
| | (0.361) | (0.778) | (0.302) | (0.247) |
| Rural | 0.330 | 3.047*** | 1.792*** | 0.904*** |
| | (0.244) | (0.551) | (0.235) | (0.225) |
| Migrant Laborer | | 0.215 | -0.181 | |
| | | (0.214) | (0.139) | |
| Caste: No Response | -0.313** | 0.391* | -0.431*** | -0.525*** |
| | (0.147) | (0.229) | (0.141) | (0.136) |
| Caste: Non-vulnerable ST | -0.219 | 0.551*** | -0.375*** | -0.603*** |
| | (0.137) | (0.177) | (0.132) | (0.127) |
| Caste: OBC | -0.129 | -0.010 | -0.127 | 0.120 |
| | (0.099) | (0.145) | (0.096) | (0.093) |
| Caste: SC | -0.264*** | 0.167 | -0.154 | -0.171* |
| | (0.100) | (0.144) | (0.099) | (0.094) |
| Caste: Vulnerable ST | -0.055 | 0.801*** | -0.571*** | -0.465*** |
| | (0.148) | (0.181) | (0.140) | (0.136) |
| Aadhaar*Female | -0.560** | 1.080** | 0.269 | 0.106 |
| | (0.274) | (0.500) | (0.288) | (0.287) |
| Aadhaar*Third Gender | -0.773** | -2.157* | -0.065 | -0.678** |
| | (0.384) | (1.270) | (0.327) | (0.281) |
| Aadhaar*Rural | 0.087 | -1.302** | -1.547*** | -1.204*** |
| | (0.254) | (0.561) | (0.246) | (0.235) |
| Constant | -1.399*** | -4.290*** | -1.441*** | -0.986*** |
| | (0.197) | (0.530) | (0.201) | (0.179) |
| Observations | 4,511 | 4,511 | 4,511 | 4,511 |
| Log Likelihood | -2,683.464 | -1,561.466 | -2,787.564 | -2,929.898 |
| Akaike Inf. Crit. | 5,392.928 | 3,150.931 | 5,603.127 | 5,885.796 |

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4: Approach 2 - Probit Regression (log odds ratio)

| | Dependent variable: | | | |
|---|---|---|---|---|
| | PDS | MGREGA | Savings Account | SIM card |
| | (1) | (2) | (3) | (4) |
| Has Aadhaar | 1.277*** | 0.605*** | 1.342*** | 1.067*** |
| | (0.109) | (0.209) | (0.111) | (0.103) |
| Female | 0.588*** | -0.548** | 0.006 | -0.218 |
| | (0.156) | (0.251) | (0.164) | (0.163) |
| Third Gender | -0.369* | -0.490 | -0.018 | 0.609*** |
| | (0.189) | (0.350) | (0.170) | (0.150) |
| Rural | 0.193 | 1.440*** | 1.081*** | 0.543*** |
| | (0.142) | (0.232) | (0.139) | (0.135) |
| Caste: No Response | | 0.145 | -0.114 | |
| | | (0.113) | (0.086) | |
| Caste: Non-vulnerable ST | -0.197** | 0.180 | -0.266*** | -0.325*** |
| | (0.088) | (0.125) | (0.086) | (0.084) |
| Caste: OBC | -0.127 | 0.297*** | -0.228*** | -0.374*** |
| | (0.082) | (0.099) | (0.081) | (0.079) |
| Caste: SC | -0.077 | -0.006 | -0.077 | 0.073 |
| | (0.059) | (0.078) | (0.058) | (0.057) |
| Caste: Vulnerable ST | -0.157*** | 0.079 | -0.093 | -0.105* |
| | (0.060) | (0.078) | (0.060) | (0.058) |
| Migrant Laborer | -0.028 | 0.441*** | -0.350*** | -0.288*** |
| | (0.089) | (0.102) | (0.086) | (0.084) |
| Aadhaar*Female | -0.336** | 0.530** | 0.137 | 0.038 |
| | (0.165) | (0.258) | (0.171) | (0.170) |
| Aadhaar*Third Gender | -0.539*** | -0.864* | -0.083 | -0.414** |
| | (0.205) | (0.483) | (0.187) | (0.170) |
| Aadhaar*Rural | 0.057 | -0.495** | -0.933*** | -0.729*** |
| | (0.149) | (0.238) | (0.145) | (0.142) |
| Constant | -0.845*** | -2.208*** | -0.877*** | -0.602*** |
| | (0.112) | (0.212) | (0.114) | (0.106) |
| Observations | 4,511 | 4,511 | 4,511 | 4,511 |
| Log Likelihood | -2,683.330 | -1,563.535 | -2,787.331 | -2,929.847 |
| Akaike Inf. Crit. | 5,392.660 | 3,155.070 | 5,602.661 | 5,885.693 |

Note: *p<0.1; **p<0.05; ***p<0.01

**Table 5: Average Marginal Effects** : PDS
Approach 1 : Logit Regression

| **PDS** | | |
|---|---|---|
| Aadhaar | 0.4124 | |
| Female | 0.09515 | |
| Third Gender | -0.3112 | |
| Rural | 0.08475 | |
| Caste: No Response | -0.06216 | |
| Caste: Non-vulnerable ST | -0.04225 | |
| Caste: OBC | -0.0245 | |
| Caste: SC | -0.05168 | |
| Caste: Vulnerable ST | -0.01034 | |

**Table 6: Average Marginal Effects** : MGNREGA
Approach 1 : Logit Regression

| **MGNREGA** | | |
|---|---|---|
| Aadhaar | 0.05832 | |
| Female | -0.01128 | |
| Third Gender | -0.1333 | |
| Rural | 0.2021 | |
| Caste: No Response | 0.04032 | |
| Caste: Non-vulnerable ST | 0.06057 | |
| Caste: OBC | -0.00203 | |
| Caste: SC | 0.01525 | |
| Caste: Vulnerable ST | 0.09459 | |
| Migrant Laborer | 0.02345 | |

**Table 7: Average Marginal Effects** : Savings Accounts
Approach 1 : Logit Regression

| **Savings Accounts** | | |
|---|---|---|
| Aadhaar | 0.3676 | |
| Female | 0.04998 | |
| Third Gender | -0.03993 | |
| Rural | 0.08244 | |
| Caste: No Response | -0.1022 | |
| Caste: Non-vulnerable ST | -0.08441 | |
| Caste: OBC | -0.03008 | |
| Caste: SC | -0.0384 | |
| Caste: Vulnerable ST | -0.1309 | |
| Migrant Laborer | -0.03841 | |

**Table 8: Average Marginal Effects** : SIM cards
Approach 1 : Logit Regression

| **SIM cards** | | |
|---|---|---|
| Aadhaar | 0.2779 | |
| Female | -0.06616 | |
| Third Gender | 0.09705 | |
| Rural | -0.04527 | |
| Caste: No Response | -0.1298 | |
| Caste: Non-vulnerable ST | -0.1445 | |
| Caste: OBC | 0.02461 | |
| Caste: SC | -0.04351 | |
| Caste: Vulnerable ST | -0.1135 | |

**Table 10: McFadden $R^2$**

| **PDS** | |
|---|---|
| Approach 1 : logit | 0.1081428 |
| Approach 1 : probit | 0.1066543 |
| Approach 2 : logit | 0.1081428 |
| Approach 2 : probit | 0.1081873 |
| **MGNREGA** | |
| Approach 1 : logit | 0.14696 |
| Approach 1 : probit | 0.1463699 |
| Approach 2 : logit | 0.1506752 |
| Approach 2 : probit | 0.1495495 |
| **Savings Accounts** | |
| Approach 1 : logit | 0.06126534 |
| Approach 1 : probit | 0.05976542 |
| Approach 2 : logit | 0.06693961 |
| Approach 2 : probit | 0.06701759 |
| **SIM cards** | |
| Approach 1 : logit | 0.04519954 |
| Approach 1 : probit | 0.0405447 |
| Approach 2 : logit | 0.04519954 |
| Approach 2 : probit | 0.04521624 |