Financial inclusion and state capacity:
A Cross-Nation Study on the Effect of Fiscal Capacity on Access to Financial Products and Services

Hanna Mullis
# Table of Contents

1. Introduction.................................................................................................................. 3

2. Theoretical framework.................................................................................................. 5
   2.1. Previous research on financial inclusion and fiscal capacity .................................. 6
   2.2. Economic development ......................................................................................... 6
   2.3. Economic formalization ....................................................................................... 9
   2.4. State capacity ..................................................................................................... 11
   2.5. Financial inclusion ............................................................................................. 11

3. Research design ............................................................................................................. 14
   3.1. Method ................................................................................................................. 14
   3.2. Data selection ...................................................................................................... 16
   3.3. Dependent and Independent variables ................................................................... 17
      3.3.1. Dependent variable: Financial inclusion ....................................................... 17
      3.3.2. Independent variable: Fiscal capacity ............................................................ 18
   3.4. Control variables ................................................................................................ 18
      3.4.1. Official development assistance as a percentage of GNI and Personal remittances, received (% of GDP) ......................................................................................... 18
      3.4.2. Human Development Index (HDI) ................................................................ 19
      3.4.3. Rule of Law ................................................................................................ 19
      3.4.4. Unemployment .............................................................................................. 20
      3.4.5. Other variables ............................................................................................ 21
   3.5. Problems and limitations ..................................................................................... 22

4. Results and analysis ...................................................................................................... 24
   4.1. Descriptive statistics ......................................................................................... 24
   4.2. Regression analysis ......................................................................................... 25
   4.3. Discussion .......................................................................................................... 32

5. Conclusion .................................................................................................................... 34

6. References .................................................................................................................... 37
1. Introduction

According to the World Bank Group (2018), almost two billion people have no access to basic 
finance, affecting their ability to save for the future and send and receive payments. This lack of 
access is not entirely voluntary and largely impacts marginalized groups in society most severely 
(World Bank, 2018), making the lack of financial inclusion a political problem. Financial 
inclusion as a topic has arisen in discourses concerning poverty reduction (Chibba, 2008; Beck et 
al., 2009; Park and Mercado, 2018) and inclusive development (Fisher and Sriram, 2002; Rojas- 
Suarez, 2010; Cull and Murdoch, 2017). Financial inclusion is defined by the World Bank as a 
phenomenon where “individuals and businesses have access to useful and affordable financial 
products and services that meet their needs… delivered in a responsible and sustainable way” 
(The World Bank, 2018). As the World Bank points out in their primer to financial inclusion, the 
ability to save, manage risks, and invest in health and education can really improve a person’s 

Improving welfare outcomes in the context of widespread poverty is a central component 
of development in theory and in practice. A classic definition of development, often utilized by 
contemporary official development assistance programs, can be found in Amartya Sen’s work, 
“Development as Freedom” (1999), where he argues that “Development consists of the removal 
of various types of unfreedoms that leave people with little choice and little opportunity of 
exercising their reasoned agency”. Within his framework of development, Sen develops the 
capability approach, where the removal of factors limiting freedom, “poverty as well as tyranny, 
poor economic opportunities as well as systematic social deprivation, neglect of public facilities 
as well as intolerance or over activity of repressive states” guarantee “the substantive freedoms 
of individuals, seen as active agents of change rather than as passive recipients of dispensed
benefits” (1999). The aim of this paper is to examine the impact of degree of state capacity on alleviating individual barriers to economic opportunities vis-a-vis financial inclusion, hence enhancing individuals’ capabilities of having agency over their own life situation. While financially inclusive policy can be of assistance in state capacity-building, some baseline of state capacity is a prerequisite for bolstering financial inclusion, and not vice-versa. The research question is thus “What effect does fiscal capacity have upon access to finance?”

The definition of state capacity is a state’s ability to carry out fundamental legal and fiscal activities, which is classified from a low to high level. Low state capacity is associated with endemic poverty, civil wars, and internal fragmentation, while high state capacity is associated with inclusive and stable state apparatuses (Johnson and Koyama, 2017). There are two primary components of state capacity, consisting of legal capacity and fiscal capacity. Legal capacity is the ability to administer justice across the state’s territory and fiscal capacity is the ability to collect tax revenue (Johnson and Koyama, 2017). This paper will focus specifically on the role of fiscal capacity.

The relationship between financial inclusion and fiscal capacity is studied by using regression analysis. The regression analyses build upon cross-sectional data that includes all regions of the world, with data from countries at different levels of economic development and financial inclusivity. A framework composed of theoretical discourses on financial inclusion, state capacity, and economic development is utilized. This framework lands in an argument focused on economic formalization, wherein economic formalization proves deleterious to barriers that weaken individuals’ capabilities to exercise economic agency and states’ capacities to govern.
The paper shows that there is a positive and statistically significant relationship between state capacity and financial inclusion. When controlling for other theoretically relevant variables, the relationship’s coefficient becomes somewhat reduced and less statistically significant. The relationship is isolated from other determinants of financial inclusion. Further research of qualitative character is needed to generate new theories on what drives financial inclusion.

The paper begins with a review of previous research, followed by a presentation of the theoretical framework, which is utilized to form the hypothesis of the presence of a positive relationship between financial inclusion and fiscal capacity. The research design is then presented, with discussions on the selection of data, operationalizations of the theoretical framework into the choice of dependent and independent variables, including the choice of control variables, followed by an examination of the problems and limitations present in the study. The result of the regression analyses is later presented and evaluated. The paper ends with some commentary on the results and suggestions for future research conclude the paper.

2. Theoretical framework

This section expands on the discussion of financial inclusion and fiscal capacity in the introduction, conveying the theoretical background to the relationship in question. Previous research on the proposed relationship is first discussed, followed by a brief breakdown of the components that underpin the theoretical arguments driving the relationship: development, economic formalization, and state capacity. This chapter is then concluded with the paper’s theoretical argument and hypothesis on the relationship between financial inclusion and fiscal capacity.
2.1. Previous research on financial inclusion and fiscal capacity

In the political science and development literature, the relationship between financial inclusion and fiscal capacity has never been outright established despite relationships between tangentially associated phenomena having been established with regards to both financial inclusion and fiscal capacity independently (see, among others, Beck et al., 2009; Besley and Persson, 2009; Chibba, 2009; Rojas-Suarez, 2010; Allen et al., 2012; Gaspar et al., 2016). Research on the subject has consequently only been carried out in relation to a handful of topics related to economic development. The literature has recently been ascribing more and more benefits to financial inclusion, including poverty reduction (see Chibba, 2008; Beck et al., 2009; Park and Mercado, 2018) and growth in consumption and household welfare (Beck et al., 2009; Seck et al., 2017; Mwangi and Ateino, 2018). Much of the previous research thus far has concentrated on the relationship between financial inclusion and an assorted group of macroeconomic development indicators (Honohan, 2008; Park and Mercado, 2018) or theoretical frameworks and case studies for developing and deepening financial inclusion in developing countries (Park and Mercado, 2015). However, a focused overview on financial inclusion and an association with state capacity is still missing.

2.2. Economic development

International development is widely seen as a desirable goal in fighting absolute poverty and raising people’s living standards. This consensus is exemplified by the Sustainable Development Goal program, and its predecessor the Millennium Development Goals, both of which were unanimously voted for in the United Nations Grand Assembly (United Nations, 2019). These commitments both attempt to view development as comprehensively as possible by tying
indicators relating to inequality, poverty, environmental degradation, and peace, among others, together.

The history of international development in theory and in practice has not always been as holistic. International development was long understood to be synonymous with linear economic growth as exemplified by absolute growth in Gross Domestic Product (GDP) (Sen, 1999), or the total value of all final goods and services produced in a country in one 12-month period. Over time the singular focus on economic growth has been challenged theoretically and empirically. In this section, a summary of the shifts in theorizing and assessing international development is presented, concluding with a theoretical framework that informs the choice of data and research question.

One of the most prominent doctrines in international development is that of modernization theory, where states follow a similar linear sequence of development paths in terms of economic evolution. Modernization is classified categorically, starting from a traditional society, where production and output was constrained by limitations in technological advances and where labor is confined to unproductive tasks. In accordance with Wilhelm Rostow’s influential five-step model of development, the final stop is an abstract society that is postindustrial “beyond consumption” that (1959). In this society, resources are fully utilized and expanding families no longer feel economic insecurity or deprived of consumption. Economic and socio-political transformation are overtly interlinked in this model, with economic transition being driven in large part by increased demand for “non-traditional” products such as financial products and imports to fit materialist norms. The primary indicators of development in this model of development are growth in GDP, capital accumulation (including human capital), and industrialization (Rostow, 1959).
In the 1990’s theories on international development shifted towards a framework of human development. Income inequalities and inabilities to “take-off” in accordance with Rostow’s model after years of development assistance to the world’s poorest countries caused the United Nations and other actors to reevaluate their approach to development programs. According to this critique, governments were not held accountable in allocating development assistance fairly and the global poor were framed as “an object of charity” rather than potential agents (Offenheiser and Holcombe, 2003). An international development paradigm that emphasizes inclusivity began to take hold, following the human rights-based approach (Cornwall and Nyamu-Musembi, 2004). Poverty reduction, human rights, inclusion (economic, political, social), and the formulation of the capability approach were synthesized into Amartya Sen’s highly influential book, “Development as Freedom”. The capability approach focuses on what individuals are able to do. In this approach, substantive freedoms such as the ability to engage in economic activity or participate in civic life is emphasized over utility maximization and access to resources (Sen, 1999). Agency is based on the capability to exercise one’s substantive freedoms (Sen, 1999) rather than the neoclassical economic definition of an agent as an actor who makes utility maximizing decisions (Levin and Milgrom, 2004).

An international development paradigm that prioritizes social inclusion has figured prominently in the international development programs of the past two decades (Gupta et al., 2015). Socially inclusive societies are safer, more stable, and meet the baseline conditions for economic transformation and growth (Abott et al., 2016). The consequences of social inclusion on economic growth and state capacity suggest a structural element to exclusionary dynamics present in many societies. Socially inclusive societies are based on “recognition of rights and responsibilities, accountability and judgement and of the fundamental equality of all and on
provision of life chances for all members of society to participate in the activities of society” (Silver, 2015). Inclusive development as such highlights inclusion of excluded people fulfillment of their capabilities (Gupta et al., 2015), in tandem with Sen’s work from 20 years ago. In this paper, attention will be paid to financial inclusion as a mode of the inclusive international paradigm.

2.3. Economic formalization

Modernization theory presumes the disappearance of “traditional” economic activity following economic transformation and growth into the second step of Rostow’s five step developmental paradigm (1959). Be that as it may, economist Hans Singer found the static nature of unemployment and underemployment during a period of robust economic growth in developing countries during the 1960’s to be baffling. He postulated a “dualism” of the labor market, where advances in healthcare meant the population was growing faster than the amount of “modern” jobs that could be created and as such were being absorbed by the “traditional” sector (Singer, 1970). Following his publication, the ILO began a multidisciplinary investigation into this employment puzzle in developing countries. Singer was a member of the first “employment mission” in Kenya that studied this phenomenon, where he noted the existence of two economic structures existing simultaneously; the industrialized, capital-intensive formal sector and the labor-intensive “traditional” sector, which came to be known as the informal sector (ILO, 1972).

The reality of dual economies with disproportionately large informal sectors have mixed, and varied, effects on a state, firms within its boundaries, and workers. For workers, the primary benefits of informal economies are the opportunities to evade taxes and potentially receive a higher pay than possible in the formal sector (Flodman Becker, 2004; Günter and Launov, 2012; Kundt, 2017; Kangrave et al., 2018). Chen (2007) notes that these effects stem primarily from
whether or not the firm is formal, meaning workers rarely have the power to make these decisions. The negatives are more numerous, including an inherent vulnerability of finding regular employment and stable incomes, lack of legal protections in terms of occupational health and safety, reduced ability to collectively bargain, and inability to register for benefits such as pensions and unemployment insurance (Flodman Becker, 2004; Chen, 2007; ILO, 2007; Günter and Launov, 2012; ILO, 2018; WIEGO, 2019). Firms have the benefit of evading taxes and cumbersome regulations, such as labor codes (de Soto, 2003) but have a low potential for growth in the broader market due to high barriers of entry and invisibility that evades formal investment (Flodman Becker, 2004; Chen, 2007). States gain a sector in which surplus labor flows into, putting public resources under less strain, but as a whole lose out on the existence of informal economies. The informal economy is generally unresponsive to policy changes due to incompetence and corruption, but this effect is diminished in the presence of effective governance structures (Guha-Khasnobis et al., 2005). The evasive character of much of the informal economy makes it difficult for states to monitor, thus hampering observed growth to a mere share of total output potential (Flodman Becker, 2004; ILO, 2009).

Campaigns to promote formalization have been promoted by the ILO along with other labor interest groups. While a total eradication of the informal sector is not necessarily the final goal, the myriad benefits of formalization to the laborer, the firm, and the state have been observed by many. Laborers would receive more legal recognition, and thus protection in areas of occupational health and safety, unionizing, health insurances, and saving for retirement (Flodman Becker, 2004; ILO, 2007; WIEGO, 2019). Some argue further that formalization establishes a social contract between the marginalized and the state as paying taxes is assumed to increase government accountability (Bates and Lien, 1985; Bräutigam, 2002). Firms stand to
gain more in judicial terms with enforceable contracts, stronger property rights, access to public infrastructure, broader finance and market information, and clear bankruptcy or liquidation rules (Flodman Becker, 2004; WIEGO, 2019). Within the context of the capability approach, increasing formalization increases the substantive freedoms of many of the stakeholders in a national economy.

2.4. State capacity

State capacity is an ideal-type in which typologies of low and high state capacity reflect the ability of a state to enforce its rules across the entirety of its territory, e.g. legal capacity, and its ability to garner enough tax revenue from the economic activities within its territory to fund the implementation of its policies, e.g. fiscal capacity (Johnson and Koyama, 2017). In practice, low state capacity is associated with endemic poverty, civil wars, and internal fragmentation, while high state capacity is associated with inclusive and stable state apparatuses (Johnson and Koyama, 2017). Fiscal and legal capacity are related, as tax revenues fund the ability of an administrative body to implement its policies and policies enable the administrative body to redistribute resources and provide public goods (Bräutigam, 2002; Besley and Persson, 2009).

The paper prioritizes fiscal capacity, as it ties in better to an analysis of financial inclusion. Financially inclusive societies better enable states to monitor economic activity, enabling the state to broaden their tax base (Besley and Persson, 2014), giving states an incentive to promote financial inclusion in their own development processes.

2.5. Financial inclusion

Financial inclusion as a method to fight absolute poverty is a part of an inclusive development framework. According to the authors of a report on the Global Findex database where the data on financial inclusion is taken from, inclusive financial systems provide an infrastructure “allowing
broad access to financial services, without price or nonprice barriers to their use” (Demirguc-Kunt and Klapper, 2012). The World Bank includes the availability of financial services in their definition of financial inclusion (World Bank, 2018). Given the size and scope of their global database, much of the literature on financial inclusion follows those definitions.

In the Findex database, financial inclusion is classified into three distinct categories: access, ease of use, and availability. This paper chooses to only analyze access as the operationalization of financial inclusion, as it is the most obvious way to observe whether or not individuals and small to medium sized enterprises (SME’s) have access to a broader country-wide financial infrastructure. Honohan (2008) writes that “the main focus of current policy concern with regard to household finance in developing countries is the perceived need to increase the access of poor households to basic financial services: deposits, payments, insurance and credit”. This is more or less the same case for SME’s (see IMF, 2019). Ease of usage and geographic availability of these services inform crude access to a bank account, not the other way around.

Financial inclusion came to prominence in the development discourse as a result of the successes and failures of the microcredit phenomenon. The microcredit industry located a market failure, in which more entrepreneurs than anticipated requested, and appropriately utilized, credit for investing in small level firms, with high levels of payback to the creditors (Mosley and Hulme, 1998). However, the microcredit industry is not well-regulated, with some customers having been unable to repay their loans with high interest fees attached (Cull and Murdoch, 2017). Likewise, the microcredit movement historically did not offer full financial products and services (Fisher and Sriram, 2002; Cull and Murdoch, 2017), especially as defined by the parameters of the Global Findex database. By financial products and services, the World Bank is
specifying means for transactions, payments, savings, credit, and insurance that are universally accessible (Demiguc-Kunt and Klapper, 2012).

By extension, financial inclusion offers the benefits of the microcredit movement alongside opportunities to make monetary transactions and save for the future. Savings in particular are associated with the ability to smooth consumption over time, which means having emergency funds on hand given an exogenous shock to one’s income; ability to invest in social programs like education for the children of a household, which has positive long-term benefits for the household as a whole; and give financial independence to individuals, which is associated with empowerment of marginalized groups in society, particularly women (World Bank, 2018).

There are seven primary barriers to an expansion in financial inclusion: a large informal sector; “worry-free” attitudes that affect savings, spending, and borrowing patterns; poor access to formal credit; absence or low presence of formal financial institutions; market concentrations (i.e. competition in different market sectors); weak enabling environment for private businesses (e.g. corruption); and poor governance in areas such as monetary policy, land ownership, and regulatory frameworks (Chibba, 2009). This list of barriers demonstrates that there is a clear theoretical relationship between financial inclusion and economic formalization, and in turn, state capacity. In essence, a reduction in size of the informal economy, a larger penetration and deepening of formal financial institutions, and an improved legal capacity should theoretically increase the share of a population that has access to financial products and services.

In concluding this chapter, there is a theoretical case to be made for an empirical relationship between financial inclusion and state capacity. Every country possesses some degree of state capacity, but what degree is meaningful in enabling access to formal financial institutions? In the following section, a statistical analysis of the operationalizations of these
concepts and controls for the possible relationship will be made to ascertain if a higher degree of state capacity financial inclusion positively affects access to finance.

3. Research design

In this section, the research design of the paper is presented. The choice of method introduces the section, which is then followed by a theoretical discussion on the selected indicators. There, the dependent, independent, and control variables will be introduced and described. Lastly, problems and limitations of the chosen method, the data, and overall research design will be evaluated.

3.1. Method

Regression analysis is applied in investigating the relationship between financial inclusion and fiscal capacity as it can provide evidence of covariance between two variables, given the ability to isolate the relationship from other explanatory factors (Stock and Watson 2015). A simple linear regression with the independent and dependent variables is first constructed. The model comprises of the dependent access to a bank account (\(\text{account}\)) and the independent percentage of tax revenue a country receives as a percentage of GDP (\(\text{taxrev}\)). Subscript \(i\) denotes the specific country and \(\epsilon\) is the error term representing the margin of error where the observed data deviates from the model’s estimated effect.

\[
\text{account}_i = \beta_0 + \beta_1 \text{taxrev}_i + \epsilon_i
\]

Simple bivariate regressions have several weaknesses in evaluating any relationship between the variables. These weaknesses are discussed later in this section. Additional multivariate regressions, containing various control variables, will be conducted to isolate the relationship between financial inclusion and fiscal capacity. An ordinary least squares (OLS)
regression model is employed in hypothesis testing in order to minimize the sum of squares in the difference between the observed values of the dependent variable and the predicted values from the linear function (Stock and Watson, 2015). In short, the smaller the differences, the better the model fits the data.

Multivariate regression is useful in estimating the effect of the independent variable on the dependent variable, *ceteris paribus*, considering the possibility to control for the effect of other factors that could have an impact on the independent variable (Stock and Watson, 2015). The full multivariate model, incorporating controls, is the following:

\[
account_i = \beta_0 + \beta_1 taxrev_i + \beta_2 ODA_i + \beta_3 remit_i + \beta_4 HDI_i + \beta_5 rol_i + \beta_6 unemp_i + \varepsilon_i
\]

This model builds upon the initial bivariate model and also includes official development assistance, remittances, the Human Development Index, rule of law, and unemployment rates as controls. These choices are motivated in section 3.4.

The data scrutinized in this paper primarily originate from surveys using random sampling by each respective countries’ economic ministries. The amount of observations is large enough to assume normal distribution, despite random sampling, given the central limit theorem (Stock and Watson, 2015). The assumption of normal distribution means the sample mean should also lie close to the population mean, which is important for the study’s generalizability (Stock and Watson, 2015). Given that the data being analyzed is non-experimental, multivariate regressions allow the analysis to hold the control variables fixed to chart the effects of the primary dependent variable on the independent variable. Although experimental settings offer the best foundations for minimizing internal validity problems, multivariate regressions do offer an ability to at least control for other related estimators in determining causality.
The analysis is conducted using cross-sectional data on financial inclusion from 95 countries from all continents and of all levels of economic development. Cross-sectional data is useful for its strengths in investigating relationships involving many different variables on country-level scales (Stock and Watson, 2015). There is a lack of comprehensive longitudinal data on financial inclusion that is needed to draw any meaningful conclusions on trends in financial inclusion and development which is why OLS multivariate regressions are used in this paper as in much of the previously cited literature (see, among others, Honohan 2008; Park and Mercado 2018). The primary methodological differences in this study are the introduction of fiscal capacity into understanding the rate of financial inclusion as well as a different makeup of control variables.

3.2. Data selection

The empirical data that is investigated in this paper is drawn from a variety of international sources. The data on financial inclusion is collected from the World Bank’s Global Findex database, which in turn is from over 150,000 nationally representative survey answers from over 140 economies (World Bank, 2017a). The data on tax revenue is also from the World Bank’s data library. This data has been collected by the International Monetary Fund, Government Finance Statistics Yearbooks and data files, and World Bank and OECD GDP estimates (World Bank, 2017a).

The control variables were by and large also from World Bank data sources, with the exception of the observations on the Human Development Index (HDI) and the unemployment rate, which comes from the United Nations Development Program (UNDP) and the International Labor Organization (ILO), respectively. The HDI is a composite index which sources its data from the UNDP, ILO, World Bank, and World Health Organization (WHO). Unemployment is
the calculated share of the total labor force without work, from a modeled ILO estimate. The official development assistance as a percentage of GNI was collected from the World Bank’s data library, based in turn on a pre-existing OECD database. The Rule of Law Index is collected by a project supported by the World Bank called the Worldwide Governance Index, a “research dataset summarizing the views on the quality of governance provided by a large number of enterprise, citizen and expert survey respondents in industrial and developing countries” collected from “a number of survey institutes, think tanks, non-governmental organizations, international organizations, and private sector firms” (World Bank, 2017b). All data is from the year 2017 in order to try securing the unbiasedness of the estimators and, by extension, the external validity of the study given missing data from many of the entities that were removed from the final dataset.

3.3. Dependent and Independent variables

3.3.1. Dependent variable: Financial inclusion

Financial inclusion as a concept is defined by the World Bank as “individuals and businesses have access to useful and affordable financial products and services that meet their needs – transactions, payments, savings, credit and insurance” (World Bank, 2018). Financial inclusion is comprised of variables measuring crude access, ease of access, and availability of services. In this paper, financial inclusion is operationalized into a variable that measures the percentage of a country’s population over the age of 15 that has access to a bank account. There are limitations to this operationalization: it does not take enterprise into account, nor does it take the availability of the services in relation to the needs of businesses and individuals being provided into account. Absolute access to a bank account for those over the age of 15, however, best reflects the penetration of financial products as explained in section 2.5. It also best illustrates possible
structural barriers at a cross-country level that individuals and businesses face in using financial products, such as cost, information asymmetries, and availability concerns (Demirguc-Kunt et al., 2015).

3.3.2. Independent variable: Fiscal capacity

Fiscal capacity is the ability of a government to garner enough tax revenue from the economic activities within its territory to fund the implementation of its policies (Johnson and Koyama, 2017). Fiscal capacity in this paper is operationalized as tax revenues as a percentage of GDP. There is some variability in this measure, as some low-income countries have relatively high tax-to-GDP ratios due to the taxation of resources while other middle-income countries can have lower tax-to-GDP ratios, which reflects upon individual tax policy decisions (Zee, 1996). Nonetheless, the measure still best represents the ability to administer taxes and policy over its territory and is therefore a good representation of fiscal capacity.

3.4. Control variables

3.4.1. Official development assistance as a percentage of GNI and Personal remittances, received (% of GDP)

The Organization of Economic Co-operation and Development (OECD) defines official development assistance (ODA) as “government aid designed to promote the economic development and welfare of developing countries” (OECD, 2019). This is chosen as a control variable as the attainment of development aid has a well-researched effect on development. Much of the literature shows a positive relationship economic growth (Addison et al., 2005; Reddy and Miniou, 2009), with some criticizing the insinuation of direct causality between foreign aid and economic development as spurious (Svensson, 1999; Burnside and Dollar, 2004). Some argue that foreign aid should be more targeted towards directly helping the poor (Klasen,
2003) and/or building state capacity (Kenny, 2006). Research on instances of targeted poverty relief have shown positive effects (Gomanee et al., 2005).

Another variable similar to ODA are personal remittances received as a percentage of GDP. Remittances are sent to individuals from kin working abroad and can be used as a supplemental income, a form of insurance, and a way to invest locally in entrepreneurial activity (Orrenius and Zavodny, 2010). Remittances can easily be received electronically, which is theoretically interesting in studying the growth of financially inclusive institutions and how this affects households (Toxopeus and Lensink, 2007).

3.4.2. Human Development Index (HDI)

HDI is a statistic composite index consisting of life expectancy, education, and per capita income. These indicators rank countries into tiers of human development, an approach that centers people’s wellbeing, their opportunities to develop and use their capabilities, and the freedom to make their own choices (UNDP, 2019). The index was developed in response to Sen’s capability approach, attempting to measure the capabilities of individuals around the world to utilize substantive freedoms. This data is collected from a pre-existing database from the United Nations Development Programme (UNDP), with values from 2017. As a variable, it is on a discrete ordinal scale from 0 to 1, with 1 being the highest score.

3.4.3. Rule of Law

The Rule of Law Index estimates governance from a scale of -2.5 (weak) to 2.5 (strong), summarizing the perceptions of the extent to which agents have confidence in and abide by the rules of society, in particular the quality of contract enforcement and property rights, presence of non-corrupt police and courts, and the likelihood of crime and violence (WGI, 2017). Rule of law can have an impact on the propensity for an individual to open a bank account; the Swedish
ODA department Sida (2004) contends that one barrier to financial inclusion is a distrust in the soundness of financial institutions. The rule of law is roughly equivalent to legal capacity, the ability for countries to enact policies throughout their territories. While state capacity is comprised of fiscal and legal capacity, fiscal capacity and legal capacity do not have a linear relationship between themselves and should thus not present a multicollinearity concern.

3.4.4. Unemployment

Unemployment is measured in percentage terms as a share of total workforce. This variable examines labor market utilization, which is useful in determining the extent of income-generating activities. Collateral is usually required for opening a bank account (Allen et al., 2012). The most common source of income is employment earnings which, Chen (2007) notes, is often too low in the informal sector to qualify for the lowest income tax bracket. Interpreting unemployment rates in this cross-country dataset is somewhat contradictory. Johnson and Koyama (2017) write that in countries with high state capacity, high unemployment is linked to a weak informal economy. In countries with dual economies, labor surplus is generally absorbed by the informal economy, lowering the official unemployment rate and exacerbating the problems noted in section 2.3.

Data on the degree of people employed in the formal sector as opposed to the informal sector is not public, but much can be gleaned from an ILO report on the subject which verifies the tenacity of informal market sectors (ILO, 2018). The way this variable is measured, by calculating the percentage of those employed from the entire labor force, is problematic due to the paradoxical connotations of unemployment depending on the labor market structure. There is not much data stratifying employment into formal and informal sectoral categories which means the variable cannot be used to appraise the ratio of the formal economy to the informal within a
country. However, as noted above, income generated by employment is necessary for those who benefit from financial inclusion to access a bank account, so this variable will still be included in the multivariate model specifications despite its weaknesses.

3.4.5. Other variables

Some theoretically relevant control variables were left out of the multivariate model due to multicollinearity concerns or missing data from 2017. For instance, income inequality is relevant in researching poverty and social marginalization (Kanbur, 2008). However, there are many gaps in the data, with many countries not reporting at all how income is allocated within their borders. Likewise, many of the observations that are available are from many different time periods. The analysis uses data only from 2017 in order to maintain time congruity; using data from several different years for different countries represents an internal validity problem.

Another variable that could be of interest in this study is the status of civil conflicts in a country. Civil conflicts shorten individuals’ time horizon due to their violence and uncertainties, making borrowing and saving money unfeasible (Justino, 2009). This variable is to a large extent represented by the strength of rule of law. The rule of law is strongly correlated to reductions in internal conflicts and instability in governance; correspondingly, conflicts often emerge in the breakdown of the rule of law (Justino, 2009).

One last variable worth mentioning is degree of urbanization in a society. Literature on financial inclusion shows that availability of services is important in attaining access (Allen et al., 2012); likewise, state capacity is easier to administer in countries with centralized population centers (Brunt and García-Peñalosa, 2015). The possible relationship between fiscal capacity and urbanization rate presents a multicollinearity concern. Despite its relevance to the research question, it is excluded from the final model so as to reduce the possibility of multicollinearity.
3.5. Problems and limitations

Simple linear regressions are unlikely to generate consistent estimations as a result of omitted variable bias (OVB). OVB appears if financial capacity is correlated with one of the omitted variables, violating the first square assumption, or if the omitted variable is a determinant of financial inclusion (Stock and Watson, 2015). OVB leads to the dependent variable becoming biased, threatening the internal validity of the analysis. To mitigate the consequences of OVB, variables that intuitively correlate with fiscal capacity and are determinants of financial inclusion in accordance with the paper’s theoretical framework will be individually included in a multivariate regression model to control for a correlation between the independent variable and the error term. This procedure allows for conditional mean independence, a statistical concept which implies uncorrelatedness (Stock and Watson, 2015). A linear relationship between the independent and dependent variables is investigated, but that does not dismiss the potential of a nonlinear relationship.

Problems can also arise from multiple linear regressions as an analytical method. Multicollinearity, high correlation between the independent variable and another regressor, can lead to an imprecise estimation of the independent estimator (Stock and Watson, 2015). This issue can be reduced by excluding correlated variables from the regression. However, imperfect multicollinearity can remain, making the OLS estimator less precise. None of the included regressors are believed to cause perfect multicollinearity, but imperfect multicollinearity is apparent in how the modifications of the original bivariate regression affect the other control variables.

The distribution of country observations (N=95) is also problematic with regards to income level as determined by the World Bank’s income classification system and with regards
to geographic spread. 144 economies are represented in the Global Findex database and are
roughly distributed as follows: 28% of countries are found on the African continent, 28% in
Asia, 28% in Europe, 14% in Latin America and the Caribbean, 1% in Australia/Oceania, and
1% in North America. In the dataset created for this paper, roughly 26% of the observations are
from Asian countries, 40% from Europe, 19% from Africa, 2% from Australia/Oceania, 13%
from Latin America and the Caribbean along with the two countries making up North America.
Based on the World Bank income classification from 2018, none of the observations are from
low-income countries. The majority are in fact from high-income countries, with a substantial
minority of middle-income (lower-middle and upper-middle combined) countries observed in the
dataset. This does not reflect the real-world distribution, and as such is a validity problem.
However, since the World Bank updated their income classification guidelines, many countries
have become high-income countries without necessarily being regarded as economically
developed by OECD metrics, such as Brazil and the Russian Federation. In this case a measure
of the distribution of income within a country could have been interesting, as GNI per capita is a
mean value that can be skewed by outlier high factor incomes that may not reflect accurately on
inhabitants’ living standards.

Lastly, some of the data used in the dataset is represented as a share of GDP or a share of
GNI. The measures are similar, with GDP representing the sum total of all final produced goods
and services and GNI building upon GDP while also adjusting for flows of capital in and out of a
country. For most countries, there is little difference between the two sums (Boland, 2017). Due
to time constraints, this measure is not standardized through transforming all GNI values to GDP
values. This can affect the biasedness of the estimators to a certain extent.
4. Results and analysis

The results of the study will be presented and analyzed in this section. Descriptive statistics will be presented first, providing a basis for interpreting the regression analyses that follow. Bivariate regressions will be presented first, followed by broader multivariate models. A discussion of the results close the section.

4.1. Descriptive statistics

Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>95</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Account</td>
<td>95</td>
<td>68.22</td>
<td>25.34</td>
<td>14.89</td>
<td>99.92</td>
</tr>
<tr>
<td>Tax revenue</td>
<td>95</td>
<td>17.64</td>
<td>5.92</td>
<td>.07</td>
<td>33.32</td>
</tr>
<tr>
<td>ODA</td>
<td>95</td>
<td>2.02</td>
<td>4.22</td>
<td>-.48</td>
<td>24.6</td>
</tr>
<tr>
<td>Remittances</td>
<td>92</td>
<td>3.99</td>
<td>5.95</td>
<td>.023</td>
<td>32.27</td>
</tr>
<tr>
<td>HDI</td>
<td>95</td>
<td>.76</td>
<td>.15</td>
<td>.42</td>
<td>.95</td>
</tr>
<tr>
<td>Law</td>
<td>95</td>
<td>.28</td>
<td>.97</td>
<td>-1.57</td>
<td>2.03</td>
</tr>
<tr>
<td>Unemployment</td>
<td>95</td>
<td>7.35</td>
<td>5.84</td>
<td>.6</td>
<td>27.44</td>
</tr>
</tbody>
</table>

Table 1 summarizes relevant descriptive statistics of each variable, rounded to the second decimal place. Afghanistan has the lowest percentage of access to a bank account, at about 15% of those over the age of 15 having access. Denmark’s population over the age of 15 has 99.9% access to a bank account, with several other countries also having more than 95% access to a bank account. Notably, the account variable has a large standard deviation of 25.34, indicating a lot of variation from the sample mean. There are 95 observations, curated to include only data from 2017. There are only 92 observations for remittances as a percentage of GDP; when remittances are included in a bivariate model, the statistical software used for this paper, STATA, automatically drops the 3 countries missing data on remittances from the model (Singapore, United Arab Emirates, and Uzbekistan).
4.2. Regression analysis

Table 2: Multivariate linear regression results table

<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tax Revenue</strong></td>
<td>1.659015</td>
<td>0.6372896</td>
<td>1.47219</td>
<td>0.6670806</td>
<td>0.6422154</td>
</tr>
<tr>
<td></td>
<td>(4.06)***</td>
<td>(2.49)**</td>
<td>(3.96)***</td>
<td>(2.83)***</td>
<td>(2.72)***</td>
</tr>
<tr>
<td><strong>ODA</strong></td>
<td>-0.4761698</td>
<td>-2.99609</td>
<td>-0.6483761</td>
<td>-0.4751399</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.18)</td>
<td>(-6.11)***</td>
<td>(-2.85)***</td>
<td>(-2.74)***</td>
<td></td>
</tr>
<tr>
<td><strong>Remittances</strong></td>
<td>-0.6247295</td>
<td>-0.6483761</td>
<td>-0.6247114</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.72)***</td>
<td>(-2.85)***</td>
<td>(-2.74)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HDI</strong></td>
<td>48.0212</td>
<td>61.16547</td>
<td>48.15614</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.78)***</td>
<td>(4.72)***</td>
<td>(2.84)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rule of Law</strong></td>
<td>12.62416</td>
<td>11.83708</td>
<td>12.596</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.16)***</td>
<td>(5.17)***</td>
<td>(5.31)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unemployment</strong></td>
<td>-0.0113075</td>
<td>-0.4189785</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(-1.13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>38.956</td>
<td>20.04434</td>
<td>51.38632</td>
<td>8.930838</td>
<td>19.94358</td>
</tr>
<tr>
<td></td>
<td>(5.12)***</td>
<td>(1.40)</td>
<td>(7.62)***</td>
<td>(0.84)</td>
<td>(1.41)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>95</td>
<td>92</td>
<td>95</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td><strong>R^2</strong></td>
<td>0.1503</td>
<td>0.8160</td>
<td>0.1605</td>
<td>0.8130</td>
<td>0.8160</td>
</tr>
<tr>
<td><strong>Adjusted R^2</strong></td>
<td>0.1411</td>
<td>0.8030</td>
<td>0.1423</td>
<td>0.8044</td>
<td>0.8053</td>
</tr>
</tbody>
</table>

Notes: Ordinary least square regressions. Figures are coefficients with the t-value in parentheses. Significance: *** p<0.01; ** p<0.05; * p<0.10

In model (1), there is a statistically significant positive correlation between tax revenue and the access to bank accounts. This means with every unit increase in fiscal capacity, the rate of access to a bank account increases by 1.659015%. About 14% of the variance in access to a bank account is explained by tax revenue as a percentage of GDP. This result is consistent with what was theorized in section 2.5. A visualization of the relationship with each country named can be found in figure 1.
In model (2), tax revenue is controlled for ODA, remittances, HDI, rule of law, and unemployment. The primary independent coefficient drops in value but retains its statistical significance. The adjusted $R^2$ swells to 79.58% of the variance in access to bank accounts explained by the model, from just 14.11% in model (1). Such a high $R^2$ is not consistent with results from previous research. This points to an issue in the unbiasedness of the estimators. Another surprising result is the slightly negative effect of remittances upon access to a bank account. This is not consistent with the theoretical justification as elaborated upon in section 3.4.1.

In model (3), unemployment and ODA are tested to check for a possible suppressed relationship as it tests insignificant in the full multivariate model, model (2). Tax revenue increases back to roughly its original coefficient. Unemployment remains statistically
inconsequential. ODA becomes statistically significant at the 99% confidence level. The adjusted $R^2$ shows that 16% of the variance explained by the model.

In model (4), the control variables that proved most statistically significant in model (2) are tested for. The HDI coefficient increases and remains statistically significant at the 99% confidence level, while the rule of law has a similar coefficient value as with model (2). Tax revenue retains a similar result to model (2). Remittances have a similar coefficient value as in model (2). That the coefficients for the rule of law and remittances remain roughly the same point to their being mechanisms in the chain of causality between financial inclusion and fiscal capacity. The adjusted $R^2$ is roughly the same as in model (2). The steep increase between models (2) and (3) is problematic in ensuring the consistency of the estimators. Further discussion on this result can be found in section 4.3.

In model (5), all variables but unemployment are included. The results are largely the same as in model (2), the full multivariate model. From this result, it can be concluded that unemployment, or at least how it is operationalized in this paper, is not a determinant in analyzing the relationship between financial inclusion and fiscal capacity. Likewise, it seems that the relationship between financial inclusion and fiscal capacity is indeed affected by an intermediary variable, or variables, which in this case seems to be comprised of the HDI value and the Rule of Law index.

The hypothesis is supported by all models, with the independent variable being slightly less significant, at the 95% confidence level, in model (2). Tax revenue, remittances, HDI, and rule of law remain statistically significant in every model, with all except for tax revenue remaining significant at the 99% confidence level in each regression. Remittances have a negative effect in each of the models, which is not consistent with the theory in section 3.4.1.
One striking result of the multiple regression analysis is the large upswing in adjusted $R^2$ whenever remittances, HDI, and the rule of law index are included in the model. When controlling for the individual parts of the HDI composite, the adjusted $R^2$ value remains very high. This could mean that the model fits very well, suffers from omitted variable bias, is in fact a non-linear relationship, or is overfitted. An overfitted model indicates a lot of residual variation in the data sample, hampering the generalizability of the model onto past and/or future years’ data. Only when removing the HDI values and rule of law is the $R^2$ in range with the results presented by previous work using OLS regressions. As HDI is a composite index, it is possible to control for each of its components: life expectancy in years, school life expectancy in years, and GNI per capita. GNI per capita is a non-linear variable and is logarithmically transformed to preserve the linear model.

The data for school life expectancy, life expectancy, and GNI per capita are collected from the World Bank database website. Life expectancy in years is the expected number of years a person will live from birth. This data is collected from the United Nations Population Division, census and other statistical reports from national statistics offices, Eurostat: Demographic Statistics, United Nations Statistical Division, U.S. Census Bureau: International Database, and Secretariat of the Pacific Community: Statistics and Demography Programme. School life expectancy is from primary to tertiary education for both sexes and is defined as the number of years a person can expect to spend within the specified level of education. School life expectancy data is collected annually by the UNESCO Institute for Statistics. Only 32

1 School life expectancy is calculated as “the sum of the age specific enrolment rates for the levels of education specified” where “the part of the enrolment that is not distributed by age is divided by the school-age population for the level of education they are enrolled in, and multiplied by the duration of that level of education” and “the result is then added to the sum of the age-specific enrolment rates.” Accessed on: http://databank.worldbank.org/data/reports.aspx?source=1159&series=SE.SCH.LIFE
observations could be located for the year 2017, which is problematic in that it is not large enough to assume normal distribution in accordance with the central limit theorem and thus may lead to biased estimators. GNI per capita is measured in current US Dollars and is collected from World Bank national accounts data and OECD National Accounts data files. An explanation on how GNI is measured can be found in section 3.5.

**Table 3:** Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life expectancy</td>
<td>95</td>
<td>74.473</td>
<td>7.307</td>
<td>54.102</td>
<td>84.099</td>
</tr>
<tr>
<td>School life exp.</td>
<td>32</td>
<td>13.805</td>
<td>2.910</td>
<td>8.906</td>
<td>22.104</td>
</tr>
<tr>
<td>(log)GNI/capita</td>
<td>95</td>
<td>8.945</td>
<td>1.443</td>
<td>5.829</td>
<td>11.276</td>
</tr>
</tbody>
</table>

In table 3, descriptive statistics of the new variables are included. The values are rounded to the second decimal place as in table 1. While taking the natural log of GNI per capita is interesting in regression analysis as it is easier intuitively to understanding the effect of a percentage change in GNI per capita on the dependent variable, it does not say much intuitively in the descriptive statistics as to the variance of observations in the sample.

The results of new multivariate models disaggregating the HDI composite index are presented below in table 4.
### Table 4: Multivariate linear regression results table

<table>
<thead>
<tr>
<th>Model</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tax Revenue</strong></td>
<td>0.6585886</td>
<td>1.401003</td>
<td>0.2917186</td>
<td>0.4239615</td>
<td>0.4239615</td>
</tr>
<tr>
<td></td>
<td>(1.17)</td>
<td>(2.75)***</td>
<td>(1.22)</td>
<td>(1.92)*</td>
<td>(1.92)*</td>
</tr>
<tr>
<td><strong>Remittances</strong></td>
<td>-0.2828021</td>
<td>-0.5713716</td>
<td>-0.1904739</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(-0.67)</td>
<td>(-1.30)</td>
<td>(-0.54)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Life expectancy</strong></td>
<td>-0.68866</td>
<td>-0.7454793</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(-0.62)</td>
<td>(-0.62)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>School life expectancy</strong></td>
<td>1.881163</td>
<td>2.513764</td>
<td>7.027028</td>
<td>7.027028</td>
<td>7.027028</td>
</tr>
<tr>
<td></td>
<td>(1.29)</td>
<td>(1.61)</td>
<td>(4.27)***</td>
<td>(4.27)***</td>
<td>(4.27)***</td>
</tr>
<tr>
<td><strong>(log)GNI/capita</strong></td>
<td>5.159815</td>
<td>8.863061</td>
<td>7.027028</td>
<td>7.027028</td>
<td>7.027028</td>
</tr>
<tr>
<td></td>
<td>(0.82)</td>
<td>(1.34)</td>
<td>(4.27)***</td>
<td>(4.27)***</td>
<td>(4.27)***</td>
</tr>
<tr>
<td><strong>Rule of law</strong></td>
<td>12.21742</td>
<td>22.08115</td>
<td>12.72961</td>
<td>12.72961</td>
<td>12.72961</td>
</tr>
<tr>
<td></td>
<td>(2.39)**</td>
<td>(15.08)***</td>
<td>(4.95)***</td>
<td>(4.95)***</td>
<td>(4.95)***</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>30.60036</td>
<td>-15.93469</td>
<td>56.84709</td>
<td>-5.704979</td>
<td>-5.704979</td>
</tr>
<tr>
<td></td>
<td>(0.71)</td>
<td>(-0.38)</td>
<td>(13.31)***</td>
<td>(-0.38)</td>
<td>(-0.38)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>31</td>
<td>31</td>
<td>95</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td>0.8128</td>
<td>0.7683</td>
<td>0.7554</td>
<td>0.7961</td>
<td>0.7961</td>
</tr>
<tr>
<td><strong>Adjusted R²</strong></td>
<td>0.7659</td>
<td>0.7220</td>
<td>0.7500</td>
<td>0.7894</td>
<td>0.7894</td>
</tr>
</tbody>
</table>

Notes: Ordinary least square regressions. Figures are coefficients with the t-statistic in parentheses. Significance: *** p<0.01; ** p<0.05; * p<0.10

In the models above, where HDI is disaggregated into life expectancy in years, school life expectancy in years, and (log)GNI per capita, only (log)GNI per capita is significant. In model (6), only 32 observations could be found and interestingly, only the rule of law is statistically significant. The sequence in choosing what to control variables to include in the different models is the same as with models (2), (3), and (4). First, a full multivariate model is produced. Then, all variables with statistically insignificant t-values are included in their own model to check for relationships that may be suppressed by the statistically significant control variables in the full multivariate model. The following model then includes only the statistically significant control
variables to control for a possible spurious relationship. In model (9), the controls remittances as a percentage of GDP and school life expectancy were omitted because they had fewer than 95 observations. Lastly, in model (10), only control variables that had at some point been statistically significant between model (6) and model (9) were included.

In model (6), only the rule of law coefficient was statistically significant at the 95% confidence level. The adjusted $R^2$ remains elevated, with about 77% of the variance in access to a bank account accounted for in the model. The number of observations are very low due to the variable school life expectancy. This hinders this model from having any further generalizability on the population mean.

Model (7) is pretty similar to model (6), but with rule of law omitted. The main independent variable is statistically significant at the 99% confidence level. The adjusted $R^2$ is reduced to about 72% of the variation in access to a bank account being explained by the model.

In model (8), there are 95 observations again. The adjusted $R^2$ is about 75%. Rule of law is statistically significant at the 99% confidence level, while the primary independent variable loses statistical significance.

In model (9), all variables with 95 observations are included. Life expectancy remains statistically non-significant. The primary independent variable becomes statistically significant again at the 90% confidence interval. Both rule of law and the log of GNI per capita are statistically significant at the 99% confidence interval. The adjusted $R^2$ increases slightly to around 78% of the variance in access to a bank account being explained by the model.

Model (10) is noteworthy in that none of the values change after omitting the life expectancy control. The rule of law and log of GNI per capita seem to be important explanatory variables in the model. In figure 2, pairwise scatter plots charting the relationship between the
variables can be seen. All of the explanatory variables are positively correlated with access to a bank account except for remittances which is consistent with the theoretical framework. The log of GNI per capita seems to be highly correlated with the rule of law index and HDI. This suggests multicollinearity, which can produce biased estimators. However, multicollinearity should not produce inflated adjusted $R^2$ values. In the next section, the adjusted $R^2$ values in the multivariate regressions will be discussed.

**Figure 2.** Pairwise scatter plots between statistically significant explanatory variables.

4.3. Discussion

The results presented in section 4.2 lend empirical support to a covariance between collected tax revenue as a percentage of GDP and access to a bank account for those over the age of 15. This covariance has been isolated from the most commonly used determinants in the literature as reviewed in chapter 2. A central question that remains is what direction of causality exists between financial inclusion and fiscal capacity – there is a theoretical case for access to finance
increasing the ability of a state to garner tax revenues within its territory. In line with chapter 2, it seems reasonable to assert some kind of mutual reinforcement between the two phenomena. In the introduction and the theoretical framework, it is argued that fiscal capacity exists in every country, regardless of the rate of financial inclusion. Therefore, a demonstrated ability to govern with regard to fiscal capacity exists independently of rate of financial inclusion, while financial inclusion is dependent on there being a baseline of governance capability. However, the idea of a mutually reinforcing relationship between the two variables is not empirically tested in this paper. In order to quantitatively determine the direction of causality it would be necessary to use time-series data and more theory on how formal financial institutions came to be historically. State capacity has existed since the dawn of the state, but the rate of financial inclusion has come about largely due to the success in broad-based usage of modern financial products and services in industrializing and industrialized countries. A historical mechanism from historical case studies in how financial institutions became broad-based would also be useful in tandem with more time-series data on financial penetration in determining the direction of the causal relationship, for example the lag of the effect in establishing financial institutions and then financial deepening in a society. Moreover, it would be interesting to do case studies of the countries that are found on the regression line on figure 1 to delineate how financial inclusion establishes itself in a country and how fiscal capacity plays into this establishment as well as why this establishment appears in the first place.

As noted in section 4.2, a discussion of the high adjusted $R^2$ values is necessary. They do not negate evidence of the covariance between financial inclusion and fiscal capacity, but they do present an external validity issue in the further generalizability of the model generated. The high adjusted $R^2$ suggest the presence of OVB and/or a non-linear relationship as discussed in
section 3.5. With regards to OVB, as written above more qualitative research on the emergence of formal financial institutions in research on state capacity and development could be of help to find a model with a better fit. The estimators in this model may be inaccurate and as such not generalizable to new data. It is worth mentioning that the priority of this paper is to establish covariance between financial inclusion and fiscal capacity while isolating the possible relationship from other explanatory factors. The model was successful in establishing covariance, but the impact of the estimator may be inaccurate. It seems that the rule of law and HDI, or more specifically GNI per capita, are the mechanisms which best describe the rate of financial inclusion.

The outliers as presented in figure 1 are also interesting. There is no cohesive geographic clustering – results that immediately stand out include the United Arab Emirates, Afghanistan, Ethiopia, Lesotho, Cambodia, Namibia, South Africa, and Denmark. As with the countries that find themselves on the regression line, the outliers can be theory-generating. Studies using the method of difference for these outliers could be interesting in determining counterfactual conditions needed for asserting causality.

5. Conclusion

Financial inclusion is arguably a development that has been meaningful in terms of improving individuals’ lives and capabilities to exercise their substantive freedoms. What has been less clear is how state capacity affects the rate of financial inclusion in a country at all, despite much of the literature ascribing inefficiencies in governance and inadequacies in the formalization of sectors and financial institutions as structural barriers to financial inclusion. The ambition of this paper was to find empirical support for a positive covariance between fiscal capacity and financial inclusion. With economic formalization as the theoretical mechanism between
increases in fiscal capacity and deepening financial inclusion, the regression analyses indicate that there is indeed a positive relationship between the two. Theoretically, formalization increases the capacity of a state to collect tax revenue and creates incentives for individuals to diversify their normal money habits to saving and borrowing credit. From this, one could derive that account usage and fiscal capacity have a mutually reinforcing and strengthening relationship, which complicates the direction of causality and sequential chain of events that causes financial inclusion to increase with every unit increase in fiscal capacity.

This study distinguishes itself from previous research in several ways. First, a mechanism directly between financial inclusion and state capacity is proposed. While evidence may not support the mechanism as proposed in chapter 2, other mechanisms, GNI per capita and the rule of law, have been reviewed based upon evidence from the multivariate regression analyses. Secondly, the relationship has always been hinted at in previous literature when discussing barriers to financial inclusion. This paper argues for a direct relationship between the two concepts and tests it empirically. The relationship is able to be isolated, and additionally the rule of law and GNI per capita seem to be intermediary variables that function as mechanisms in understanding the relationship. Lastly, a different composition of control variables are used in this study, drawn from previous research on financial inclusion and on state capacity and governance.

New insights into the effects of one of the barriers to financial inclusion, state capacity, are demonstrated empirically. This contributes to the literature in both fields of research. Additionally, a focused overview on economic formalization, state capacity, financial inclusion, and development in line with the capability approach has been provided. This can serve as a
complement to the existing and growing body of research into financial inclusion and development.

Some validity concerns emerged upon conducting the research. Formalization as operationalized as the unemployment rate did not pan out. The operationalization of unemployment itself was vague and did not differentiate between formal and informal employment. The study can therefore not come to the conclusion that formalization is a relevant mechanism in understanding the relationship between financial inclusion and fiscal capacity despite a justified theoretical case.

For future studies, an employment rate that differentiates between formality and informality in the labor market would be more useful in ascertaining the effect of formalization upon increases in financial inclusion. Future research could also better determine the chain of causality using time-series data and forecast the growth of financial inclusion accordingly. Moreover, the causal mechanism between financial inclusion and fiscal capacity is not thoroughly investigated. A future paper using qualitative methods could better chart the mechanism’s features. Focused case studies could also investigate more intensely the role of institutions and good governance in establishing formal financial institutions and increasing the state’s capacity to tax within its territory. It provides a ripe opportunity for understanding how countries have attempted to develop their fiscal capacities and how this may have affected the establishment of formal financial institutions in their territory.
6. References


