Policy Research Working Paper 8959

Job Accessibility and Urban Transport Connectivity

Evidence from Antananarivo, Madagascar

Atsushi Iimi

WORLD BANK GROUP
Transport and Digital Development Global Practice
August 2019

Policy Research Working Paper 8958

Abstract

In recent years, there has been renewed interest in access to jobs in relation to transport connectivity. In Sub-Saharan Africa, about 14 million working age people are added to the labor market every year. Ensuring sustained access to jobs seems to be a prerequisite for inclusive and robust economic growth. The paper examines the impact of public transit connectivity on access to jobs, especially focusing on wages. Using data from Antananarivo, Madagascar, it is shown that the wages earned by commuters are systematically higher than the wages earned by those who decided not to commute and are self-employed or engaged with family businesses around their neighborhood. Proximity to public transport, especially taxi-be, is important to promote people's access to jobs. It is also found that there is a substantial gender inequality in wages in the country: Women are more likely to use buses to commute, and yet, they earn less than men. In addition, the poor tend to benefit less from public transportation. Public bus services are affordable, however, the quality of the services may remain low.

This paper is a product of the Transport and Digital Development Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The author may be contacted at aiimi@worldbank.org.

Job Accessibility and Urban Transport Connectivity: Evidence from Antananarivo, Madagascar

Atsushi Iimi[¶]

Africa Infrastructure, Transport Global Practice
World Bank Group

Key words: Access to jobs; Urban transportation; Endogenous treatment effects model.

JEL classification: C21, C26, O18, R41

[¶] Corresponding author.

I. INTRODUCTION

1. In recent years, there has been renewed interest in access to jobs in relation to transport connectivity. In the developing world, about 13 percent of youth (ages 15 to 24) are unemployed. In particular in Sub-Saharan Africa, about 14 million working age people are added to the labor market every year (McArthur, 2014). Ensuring sustained access to jobs seems to be a prerequisite for inclusive and robust economic growth. In Africa, urban unemployment is often much higher than the national unemployment rates (**Table 1**) (Girsberger and Meango, 2017). In addition, vulnerable employment, such as informal jobs and family workers, is especially high at nearly 70 percent in Sub-Saharan Africa (ILO, 2016).

2. Access to jobs is not a new topic in the literature but has long been discussed in different but relevant contexts, such as urban growth and mobility, traffic congestion, car ownership, modal choice, residential locational choice, environmental sustainability and gender inequality in the labor market. For example, Hymel (2009), using a historic highway development plan as an instrument, shows that traffic congestion weakened subsequent employment growth in U.S. metropolitan areas. Jin and Rafferty (2017) also find a similar result particularly during the 2000s.

Table 1. Unemployment rates in West Africa

		Unemployment (% of labor force)	
Country	Capital	Capital	National
Benin	Lome	6.0	0.7
Burkina Faso	Ouagadougou	20.1	2.7
Côte d'Ivoire	Abidjan	15.5	3.8
Mali	Bamako	10.7	6.3
Niger	Niamey	22.1	1.5
Senegal	Dakar	20.0	5.7
Togo	Cotonou	9.4	1.9

Sources: Girsberger and Meango (2017); WDI.

3. In a basic monocentric urban model, all employment is assumed to be in the center of a city, called the central business district (CBD). The people's demand for travel to work,

which is often a main cause of urban traffic congestion, generally depends on the marginal transportation cost and where they are located (e.g., Muth, 1969; Mills, 1972). Thus, the residential locational choice is dependent on the housing price gradient in a city (e.g., Atack and Margo, 1998; Gedal and Ellen, 2018). Therefore, whether people commute to work is essentially determined by transport connectivity and residential land values.

4. The modal choice literature has also extensively been discussing travel decisions by commuters for a very long time. In addition to pioneer studies in this area (e.g., McFadden, 1974a, 1974b; Hensher, 1986; Hensher and Greene, 2002), for example, Bento et al. (2005) examine the commute mode choices in relation to urban structure and public transit supply in the United States, showing that poor people and women are more likely to use public transportation for commuting. In particular, bus ridership is found to increase with the supply of bus transit services and the relative spatial concentration of jobs to housing. In their analysis, notably, a choice of nonmotorized transit is included, that is, walking and bicycling. Interestingly, as jobs-housing balance increases, people are more likely to walk to work, which is often referred to as the compact city.

5. Related to the above, some studies discuss job accessibility in connection with car ownership. In particular, where population density is low, quality public transportation services are often difficult to provide, and therefore, car dependence must of necessity increase. In Barcelona and Madrid, it is found that the reduction of travel time to jobs reduces the number of second or third car owners and substantially increases the number of households without a car (Matas et al., 2009). The findings in New York by Salon (2009) look consistent: New Yorkers are sensitive to travel time and costs to decide how many vehicles they own.

6. From the gender point of view, labor economics discusses gender differences in access to jobs in the labor market. Gobillon et al. (2015) analyze micro employment data in France and find that females have significantly lower access to jobs, especially, high-paid jobs, such as executive job positions. There are two different arguments: First, as Gobillon et al.

discuss, the accessibility to jobs may be more limited to women in the labor market or job assignment framework. Second, in the transportation or mobility context, women may also have disadvantage over men: Female workers are often more likely to rely on public transportation (Bento et al., 2005) or walking (Mondschein and Taylor, 2017). In addition, public transit may not be safe for women and girls (e.g., Ceccato 2017).

7. The firms' locational choice is perhaps another side of the job accessibility issue: While employees decide where they live and how they commute to work, employers or firms also choose where they are located, possibly depending on transport connectivity. According to Sweet (2014), firms are more likely to relocate if they face large-scale traffic congestion at the regional level in the Philadelphia Metropolitan Area. In general, firms prefer to locate themselves where access to transport infrastructure is good (e.g., Cieślik and Ryan, 2004; Boudier-Bensebaa, 2005; Milner et al., 2006).

8. Finally, there are an increasing number of studies that aim to develop various indices measuring job accessibility. In recent years, various GIS-based high-resolution data have become available even at the city level, which allow to estimate all kinds of connectivity and accessibility in different ways (e.g., Cheng and Bertolini, 2013; Wang and Chen, 2015; Saghapour et al. 2016). The current paper also uses some spatial data to compute transport connectivity to the central business district (CBD) and job locations.

9. The current paper aims at examining the impact of public transit connectivity on access to jobs, particularly focusing on wages earned by those who commute to work. Given the above discussion in the literature, there are two issues that constitute this result chain: First, among others, accessibility to public transit may affect people's decision on whether or not to commute to work. Second, once people are employed and decide to commute to work, the level of wages is determined. Commuting is generally costly. Thus, holding everything else constant, commuters' wages must be higher than those that non-commuters earn.

10. The paper explores this commute-wage relationship by using data from Antananarivo, the capital of Madagascar. Despite the accelerating urbanization in Africa, currently, little evidence is available for policy makers in the region. The current paper contributes to adding new evidence to the literature. Understanding how individual households decide to commute to work is essential for urban planners and transit developers. Methodologically, the paper relies on an instrumental variable (IV) technique to deal with potential endogeneity associated with the commute decision, which is another contribution to the literature.

11. The rest of the paper is organized as follows: section II provides a brief overview of Antananarivo; section III develops our empirical strategy, and section IV explains our data; section V discusses the main results; and section VI examines the robustness of the results; and then, section VII concludes.

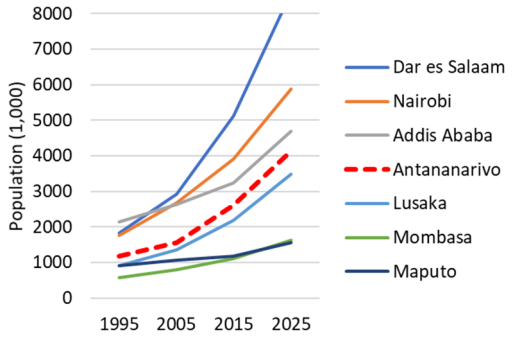
II. OVERVIEW OF URBAN TRANSPORTATION IN ANTANANARIVO

12. As experienced in many African countries, Madagascar is currently experiencing very rapid urbanization. According to UN-Habitat (2016), the population of Antananarivo, the capital of Madagascar, more than doubled in the last decade, from 1.2 million in 1995 to 2.6 million in 2015. It is the fourth largest city in East and Southern Africa, following Dar es Salaam (5.1 million), Nairobi (3.9 million) and Addis Ababa (3.2 million) (**Figure 1**). Attracting more people and more businesses, the city population is projected to continue growing to 4.1 million by 2025.

13. Similarly to other large cities in Africa, Antananarivo is the growth center in Madagascar. The country's GDP per capita was about US$401 in 2016, and the total population is about 25 million. Thus, about 10 percent of the total population lives in Greater Antananarivo. More than half of Malagasy establishments, which amount to about 120,000, are located in Antananarivo (**Figure 2**). Analamanga is a region surrounding Antananarivo and

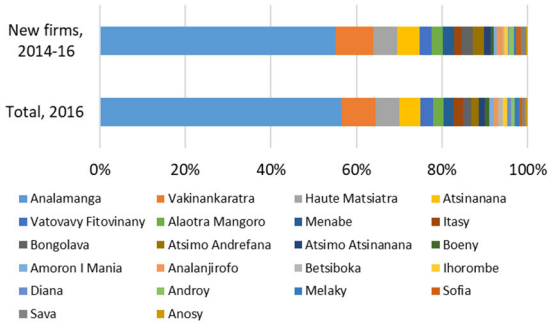
its neighboring communes. Per capita GDP in Antananarivo is estimated to US$751, which is 80 percent higher than the national level.[1]

Figure 1. Population of major cities in East and Southern Africa



Source: UN-Habitat (2016).

Figure 2. Share of registered enterprises by region



Source: INSTAT Madagascar.

14. Unfortunately, however, people do not seem to be able to take full advantage of existing job opportunities in the city, because of aggravating transport constraints. In recent years, Antananarivo has been facing unmanageable traffic congestion. Because of limited space dedicated to transport infrastructure and poorly regulated transport services, the accessible areas to the central business district (CBD) is constrained (**Figure 3**).[2]

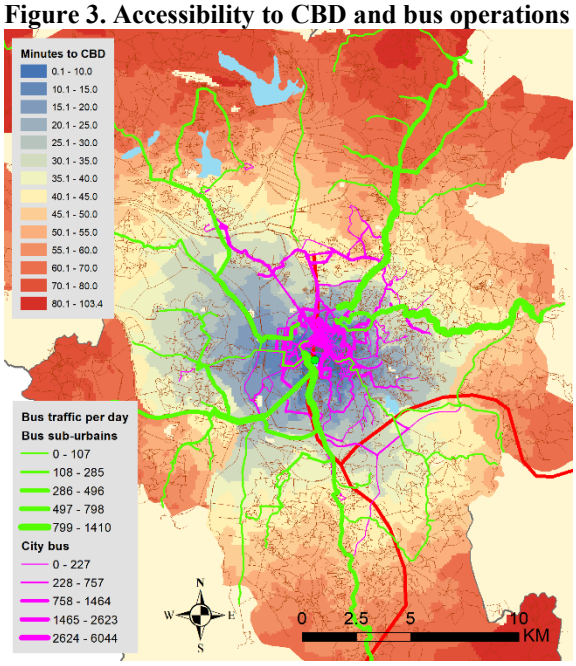
15. Since accessibility to the CBD is poor, many people prefer to live near the CBD. About 60 percent of the total city population live within 5 km of the CBD, which is assumed to be Lake Anosy in this paper (**Figure 4**). Thus, two options are left to people: The first is to pay significantly high rents and live around the CBD, and the second is to live far from the center of the city and commute to work (i.e., incurring transport costs and time).[3]

[1] According to World Bank (2017), the average household consumption per capita was about MGA 2 million in 2016. Using a share of private consumption to GDP in the national accounts data (0.79), per capita GDP is estimated at MGA 2.6 million or US$751.

[2] Global experiences indicate that many large cities use 15 to 25 percent of built-up areas for transport infrastructure (Angel et al., 2016). However, the share in Antananarivo is merely 5.7 percent.

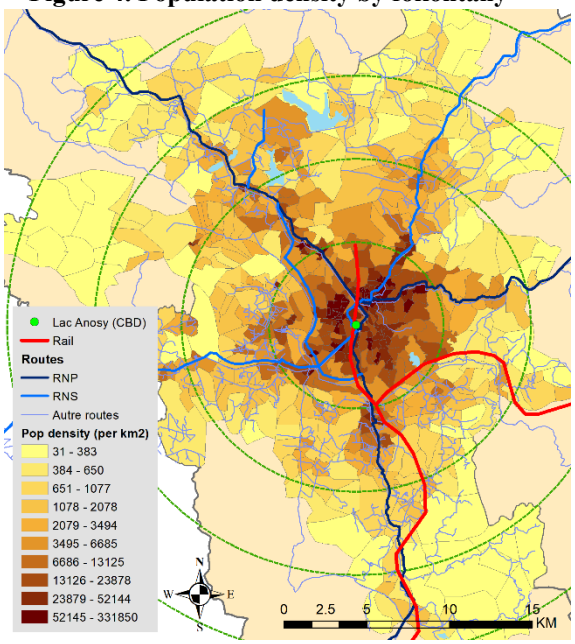
[3] The land value gradients are very steep in Antananarivo (see a companion paper, Iimi (2019)). The highest rate around Lake Anosy is about 1.1 million Malagasy ariary (MGA) or US$300 per m^2.

Figure 3. Accessibility to CBD and bus operations



Source: World Bank estimates.

Figure 4. Population density by fokontany



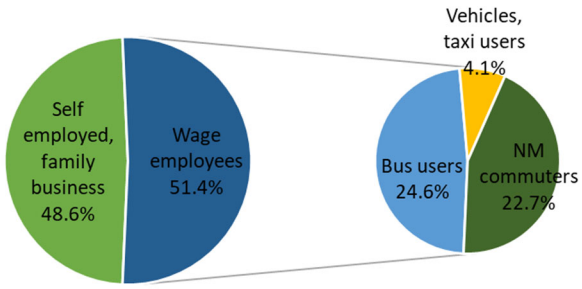
Source: Based on data by the Ministry of Health.

16. The main transport means for most of the city residents is minibus, which is called taxi-be. There are also suburban buses. It is estimated that more than 300,000 people use minibuses every day. The bus services are provided along about 130 routes by about 80 operators with over 4,000 buses. They are normally organized as cooperatives at the commune level but not well regulated or coordinated. As the result, many bus routes are duplicated and concentrated on certain routes, aggravating, not easing, traffic congestion.

17. Currently, about 25 percent of the total labor force is estimated to use buses to commute to work (**Figure 5**). According to a recent household survey in Antananarivo, in which about 2,300 households were interviewed, 4,746 of 9,350 household members were 18 to 60 years old and considered as potential labor force. Of those, 2,953 people were actually working. About half of them are self-employed or engaged with family businesses. Another half were wage employees, of which 879 persons reported how much they actually earned. Of those, about one-third, i.e., 287 persons commuted to their workplaces by bus.

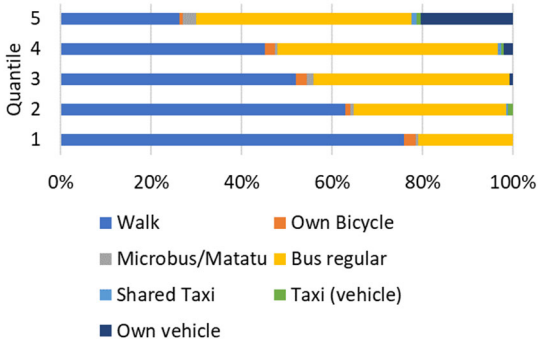
18. Poor people cannot even use public transit. The available data do not include household income but allow to estimate consumption per capita. About 75 percent of the poorest cohort walk to work (**Figure 6**). When walking is excluded, poor people are more likely to use public buses. The rich own their private vehicles or can use taxi to commute. Rich people spend more money on transportation, however, in relative terms, the burden of transportation spending is heavier to the poor (**Figure 7**). This looks to exceed the regional norm, which is less than 5 percent (Lozano-Gracia and Cheryl, 2014). Clearly, job access and transport connectivity must be of necessity related to poverty and affordability issues.

Figure 5. Wage employees and bus users in the total labor force



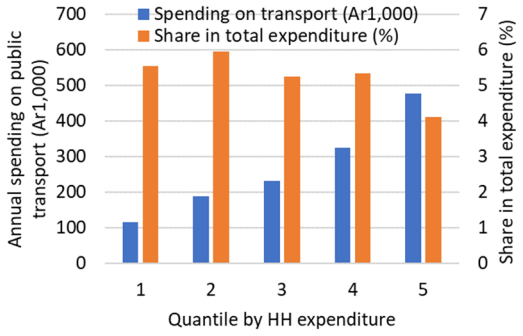
Source: World Bank estimates based on LSMS data.

Figure 6. Modal choice by consumption level



Source: World Bank estimates based on LSMS data.

Figure 7. Household spending on transport



Source: World Bank estimates based on LSMS data.

III. EMPIRICAL STRATEGY

19. To examine the impact of commuting to work on wage, as discussed above, the following wage equation is considered.

$$\ln W_i = \alpha D_i + X_i{}'\beta + \varepsilon_i \qquad (1)$$

where individual *i* is assumed to earn wage *W*, depending on her characteristics, W_i, such as age, sex and education attainment. The wage level is also dependent on whether or not her commutes to work, which is represented by a dummy variable D_i. ε is an idiosyncratic error distributed independently and identically.

20. As in the classical union-wage literature (e.g., Heckman, 1979), a problem to estimate the equation is that the decision on commuting is potentially endogenous. It is not randomly assigned. Therefore, Ordinary Least Squares (OLS) regression is likely to be biased. To deal with this issue, the following selection mechanism is considered:

$$D_i = \begin{cases} 1 & \text{if } Z_i'\gamma + X_i'\delta + u_i > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

where the dummy variable D_i is set to one if household member *i* commutes to work. The decision is assumed to depend on a set of connectivity and household characteristics, Z_i.

21. In basic monocentric urban theory, the people's demand for commuting is generally determined by marginal transportation costs and where they live. And the residential locational choice is dependent on housing price gradients (e.g., Muth, 1969; Mills, 1972; Atack and Margo, 1998). Therefore, *Z* must of necessity contain the variables representing (i) transport connectivity to workplaces, (ii) household income level, and (iii) housing conditions, such as the quality of houses, e.g., access to public infrastructure services. Of

course, the commute decision may also depend on the household member *i*'s characteristics, X_i, as well.

22. To estimate Equations (1) and (2), the paper primarily uses the endogenous treatment effects model, with the cluster-robust variance-covariance combined. Our unit of analysis is household members, not households. In each household, more than one household member may work. For obvious reasons, there would likely be common unobservables shared within a household. For endogeneity, the standard Hausman exogeneity test is performed.

IV. DATA

23. The current paper uses micro data from the Living Standards Measurement Survey (LSMS) carried out in Antananarivo in 2016. The survey interviewed about 2,300 households in 200 local areas of the city. The data cover 9,350 household members. In the sample, 2,953 people were working, of which 1,488 were wage employees. The rest were self-employed or engaged with family businesses. Since the wage information is always sensitive to those who are interviewed, only 879 household members reported their actual wages. Of those, 287 persons were commuting to work by public transit. With the observations with missing values excluded, the following analysis uses 858 observations.

24. The current paper aims to compare the level of wages between commuters and noncommuters. "Commuters" are defined by those who are employed formally (as opposed to those who work in the informal sector) and commute to their workplaces by public transportation. First, it is important to note that no information is available about wages of informal workers. Second, by this definition, those who commute by walking are not considered as commuters. Finally, those who commute by own car are not regarded as commuters, either.

25. The summary statistics are shown in (**Table 2**). The average monthly salary is about MGA 276,000 or US$78. On a simple average comparison, commuters were paid 20 percent more than noncommuters (**Figure 8**). Using geo-referenced household location data, three variables measuring transport connectivity are constructed: (i) distance to the nearest Taxi-be route (*TAXIBE*), (ii) distance to the nearest suburban bus route (*SUBBUS*), and (iii) estimated travel time to the CBD (*MINICBD*). In our sample, it takes on average 30 minutes for households to go to the CBD, with a wide variation from nearly zero minutes to 80 minutes (**Figure 9**). Also see **Figure 3** above.

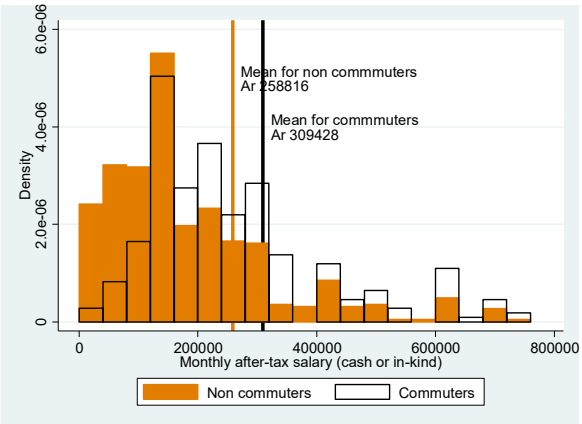
26. With respect to public transportation, many households have good access to it: The mean distance to the nearest taxi-be and suburban bus route is about 530 meters and 480 meters, respectively. Note that the distances are measured from the bus routes, not bus stops. There are no strictly designated bus stops in the Malagasy bus systems. People know where many buses stop, however, in practice, people can get on anywhere along the routes. That is why the main national roads to the center of the city, such as RN1, RN2, RN3, RN4 and RN7, are too congested currently: Buses stop and go no matter where, causing massive traffic congestion.

27. Because of our sampling strategy, several suburban areas were surveyed, in which taxi-be, which operates mainly in the center of the city, may not be available. Thus, the maximum distance from the taxi-be route is longer than that from the suburban route. Suburban bus services connect most remote areas, however, density and frequency of services are low (**Figure 10**). As the result, the poor tend to have limited connectivity to the CBD.

Table 2. Summary statistics

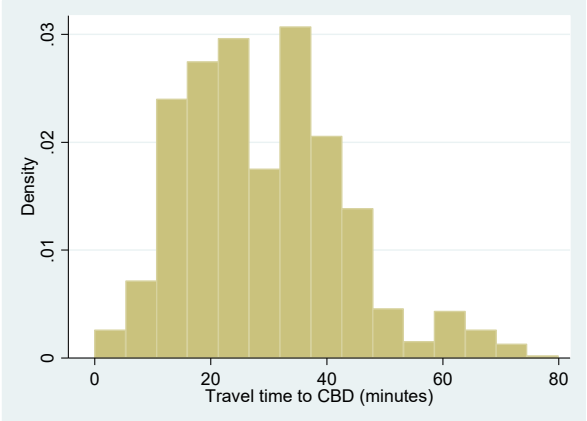
Variable	Abb.	Obs	Mean	Std. Dev.	Min	Max
Monthly salary (Ar million)	*W*	858	0.28	0.50	0.0025	10
Dummy for commuters	*D*	858	0.33	0.47	0	1
Distance to Taxibe routes (km)	*TAXIBE*	858	534.55	1067.97	0.01668	6941.74
Distance to suburban bus routes (km)	*SUBBUS*	858	484.43	437.47	0.06477	2837.85
Travel time to CBD (minutes)	*MINCBD*	858	29.57	13.69	1.290293	79.85
Housing rent paid (Ar)	*RENT*	858	28153	103038	0	2300000
Age	*AGE*	858	38.03	12.15	12	78
Sex (male=1)	*MALE*	858	0.59	0.49	0	1
Educational attainment:						
Dummy for primary education	*PRIM*	858	0.27	0.45	0	1
Dummy for secondary education	*SECO*	858	0.25	0.43	0	1
Dummy for tertiary education	*TERT*	858	0.31	0.46	0	1
Household consumption:						
Dummy for 1st quantile	*Q1*	858	0.20	0.40	0	1
Dummy for 2nd quantile	*Q2*	858	0.20	0.40	0	1
Dummy for 3rd quantile	*Q3*	858	0.22	0.41	0	1
Dummy for 4th quantile	*Q4*	858	0.21	0.41	0	1
Number of family members	*HHS*	858	4.20	1.64	1	16
Travel time to CSB (minutes)	*MINCSB*	858	3.44	2.80	0.006001	15.31
Number of rooms	*ROOM*	858	1.96	1.27	1	13
Dummy for home owner	*HOME*	858	0.40	0.49	0	1
Access to infrastructure services:						
Electricity for cooking	*POWR*	858	0.06	0.24	0	1
Private tap water	*TAPW*	858	0.14	0.35	0	1
Toilet with seat	*TOIL*	858	0.23	0.42	0	1

Figure 8. Distribution of wages by commuters and non-commuters



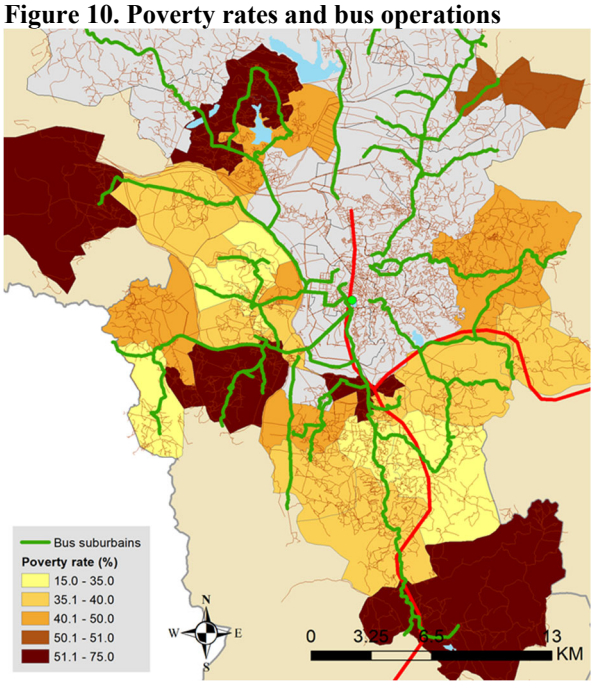
Sources: INSTAT; World Bank estimates.

Figure 9. Estimated travel time to CBD



Source: World Bank estimates.

Figure 10. Poverty rates and bus operations



Source: JICA and World Bank estimates.

V. ESTIMATION RESULTS

28. To deal with the potential selection bias, the wage equation is estimated with the endogenous treatment effects. The results are shown in **Table 3**. The first model includes the district specific fixed effects, which may capture some unobserved local characteristics. The results turned out very similar regardless of the district dummy variables are included or not.

29. Regarding the endogeneity of the commute decision, as shown at the bottom of the table, ρ is found to be significantly positive. Formally, the endogeneity test statistics are also calculated. The test statistics are estimated at 13.56 and 10.94, respectively. Both are statistically significant, suggesting that the selection of commuting is likely to be endogenous.

30. Our results indicate that commuters earned more than noncommuters. This is a main result of this paper. The coefficient of D is estimated at 0.691 or 0.674 and is statistically

significant. Thus, as expected, commuting is profitable, meaning that better employment opportunities exist away from where they live. Otherwise, people would not spend time on commuting. However, to allow people to commute to work, the transport connectivity is crucial. The commute decision is significantly dependent on the access to public transit, especially to taxi-be, minibus around the central city. The coefficient of *TAXIBE* is estimated at -0.076, which is statistically significant: Thus, holding everything else constant, those who live closer to taxi-be routes are more likely to commute. On the other hand, the proximity to suburban bus does not have a significant impact. The coefficient of travel time to the CBD is also not significant, either. Therefore, it can be concluded that the proximity to taxi-be is of particular importance to improve people's accessibility to jobs.

31. With respect to other coefficients, in the wage equation, the results imply that human capital is important to earn higher wages: The coefficient of *AGE* is positive and significant. The education variables also have all positive and significant coefficients. In particular, the impact of education is higher when the individual attains a higher level of education, such as secondary and tertiary education. For instance, *TERT* has a coefficient of 0.622, which is greater than the coefficient of *PRIM* (0.194) and *SECO* (0.284).

32. Male wages are significantly higher than female wages. The coeffect of *MALE* is estimated at 0.495, which is significant, suggesting that there is gender inequality in salary in Madagascar.

33. In the commute decision equation, it is found that younger people and women are more likely to commute by public transportation. The coefficient of *AGE* is estimated at -0.324, meaning that the elder are less likely to use buses. On the other hand, the coefficient of *MALE* has a significantly negative coefficient, indicating that women are more likely to commute by buses. These are consistent with the practical views in Madagascar.

Table 3. Endogenous treatment effects model

Dependent var.	ln*W*		*D*		ln*W*		*D*	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
D	0.691	(0.151) ***			0.674	(0.158) ***		
ln*TAXIBE*			-0.076	(0.037) **			-0.075	(0.034) **
ln*SUBUBUS*			0.027	(0.040)			0.042	(0.037)
ln*MINCBD*			0.037	(0.177)			0.049	(0.103)
ln*RENT*			0.046	(0.012) ***			0.046	(0.012) ***
ln*AGE*	0.182	(0.088) **	-0.324	(0.151) **	0.176	(0.089) **	-0.337	(0.149) **
MALE	0.495	(0.056) ***	-0.171	(0.091) *	0.493	(0.056) ***	-0.171	(0.091) *
PRIM	0.194	(0.085) **	0.283	(0.167) *	0.194	(0.085) **	0.302	(0.169) *
SECO	0.286	(0.088) ***	0.592	(0.170) ***	0.292	(0.088) ***	0.606	(0.171) ***
TERT	0.622	(0.097) ***	0.678	(0.170) ***	0.634	(0.097) ***	0.676	(0.170) ***
Q1	-0.715	(0.132) ***	-0.695	(0.215) ***	-0.694	(0.131) ***	-0.649	(0.210) ***
Q2	-0.526	(0.124) ***	-0.433	(0.189) **	-0.504	(0.122) ***	-0.359	(0.188) *
Q3	-0.530	(0.110) ***	-0.063	(0.178)	-0.501	(0.109) ***	-0.006	(0.176)
Q4	-0.386	(0.101) ***	0.022	(0.170)	-0.374	(0.102) ***	0.044	(0.170)
ln*HHS*	-0.061	(0.070)	-0.123	(0.124)	-0.062	(0.072)	-0.118	(0.119)
ln*MINCSB*	-0.011	(0.027)	0.019	(0.048)	-0.007	(0.025)	0.014	(0.047)
ln*ROOM*	0.180	(0.068) ***	0.190	(0.116) *	0.194	(0.066) ***	0.204	(0.115) *
HOME	-0.074	(0.063)	0.155	(0.132)	-0.073	(0.062)	0.162	(0.133)
POWR	0.206	(0.112) *	-0.116	(0.196)	0.199	(0.111) *	-0.100	(0.191)
TAPW	0.190	(0.105) *	-0.147	(0.191)	0.180	(0.102) *	-0.141	(0.188)
TOIL	0.050	(0.075)	-0.058	(0.134)	0.044	(0.074)	-0.067	(0.131)
Constant	10.986	(0.380) ***	0.601	(0.742)	10.970	(0.379) ***	0.375	(0.700)
Obs.	858				858			
No. of district dummies	8		8		0		0	
Wald chi2	580.29				528.90			
Inter-equation correlation parameters:								
λ	-0.426	(0.101) ***			-0.402	(0.108) ***		
ρ	0.766	(0.039) ***			0.766	(0.039) ***		
Endogenous selection test:								
Wald chi2	13.56	***			10.94	***		

34. One unexpected result may be that *RENT* has a positive, not negative, coefficient, which is statistically significant. This contradicts economic theory: In a simply monocentric urban model, there must be negative correlation between travel time and housing rents. Our results imply that those who pay more for housing are more likely to use public transit to commute to their workplaces.

35. There are several possible reasons for this. First, the current bus operations may not be provided to the poor, who may need to be served more. As discussed, the current routes may not be extended to suburban areas where many poor people live. In addition, affordability may also be an issue. The current bus fares are MGA 500 to MGA 900 or US$0.14 to US$0.24 per ride, which are not necessarily too expensive. However, taking the level of services (i.e., slow speed, too many stops, too congested and unsafe) into account, the real fares may not be so low. As the result, relatively rich people tend to benefit more from public transport services. This is consistent with the estimated household quantile effects. The dummy variables for the poor households, i.e., *Q1* or *Q2*, tend to have greater coefficients in absolute terms.

36. Second, there may be a measurement error in our rent variable. In the sample, many households actually do not pay any rent. They are homeowners and relatively rich. In the above estimation, a conventional approach to avoid the logarithm of zero was used, i.e., adding a small positive value to *RENT*. However, this may cause significant bias in the estimation results (Battese 1997). Thus, the Battese specification, which introduces dummy variables, *dRENT*, for the incidence of zero inputs, is used. The results are shown in **Table 4**. The statistical significance of *RENT* was lost, while the rest of the coefficients remain broadly unchanged: The impact of commuting on wages is significantly positive, and the proximity to taxi-be services is critical to support the commuting decision.

Table 4. Endogenous treatment effects model with Battese specification for *RENT*

Dependent var.	ln*W*			*D*		
	Coef.	Std.Err.		Coef.	Std.Err.	
D	0.691	(0.150)	***			
ln*TAXIBE*				-0.076	(0.037)	**
ln*SUBUBUS*				0.026	(0.041)	
ln*MINCBD*				0.038	(0.177)	
dRENT				0.087	(1.065)	
ln*RENT*				0.054	(0.096)	
ln*AGE*	0.183	(0.089)	**	-0.324	(0.151)	***
MALE	0.495	(0.056)	***	-0.172	(0.091)	*
PRIM	0.194	(0.085)	**	0.283	(0.167)	*
SECO	0.286	(0.088)	***	0.591	(0.170)	***
TERT	0.622	(0.097)	***	0.677	(0.171)	***

Q1	-0.715	(0.132)	***	-0.693	(0.215)	***
Q2	-0.526	(0.124)	***	-0.430	(0.191)	**
Q3	-0.530	(0.110)	***	-0.062	(0.179)	
Q4	-0.386	(0.101)	***	0.022	(0.170)	
ln*HHS*	-0.061	(0.070)		-0.123	(0.123)	
ln*MINCSB*	-0.011	(0.027)		0.019	(0.049)	
ln*ROOM*	0.180	(0.068)	***	0.189	(0.118)	
HOME	-0.074	(0.063)		0.155	(0.132)	
POWR	0.206	(0.112)	*	-0.116	(0.196)	
TAPW	0.190	(0.105)	*	-0.148	(0.192)	
TOIL	0.050	(0.075)		-0.058	(0.134)	
Constant	10.986	(0.380)	***	0.515	(1.301)	
Obs.	858					
No. of district dummies	8			8		
Wald chi2	580.96					
Inter-equation correlation parameters:						
λ	-0.426	(0.100)	***			
ρ	0.766	(0.039)	***			
Endogenous selection test:						
Wald chi2	13.77	***				

VI. ROBUSTNESS

37. Related to the selection endogeneity, one may wonder whether the instruments are statistically valid. To formally test this question, the instrumental variable (IV) regression is performed. Although this is a linear model, it is still valid to use. Not surprisingly, the results are similar to the above (**Table 5**). Commuting increases wages, and the proximity to taxi-be routes has a positive impact on the commuting choice. Human capital is important to earn more. Women are more likely to use buses to commute, and yet, they earn less than men.

The conventional endogeneity test statistics are estimated at 8.57 and 7.70, respectively. Thus, the exogeneity hypothesis can easily be rejected: The commuting selection is endogenous. On the other hand, the overidentifying restriction test statistics are 4.48 and 5.70, which cannot reject the hypothesis that there is no correlation between the selection status and the error term. Therefore, our proposed instruments are valid to control for the endogeneity of the commuting choice in the system.

Table 5. Instrumental Variable (IV) regression results

Dependent var.	ln*W*		*D*		ln*W*		*D*	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
D	1.469	(0.568) ***			1.282	(0.491) ***		
ln*TAXIBE*			-0.026	(0.013) **			-0.028	(0.012) **
ln*SUBUBUS*			0.008	(0.014)			0.015	(0.013)
ln*MINCBD*			0.036	(0.061)			0.015	(0.037)
ln*RENT*			0.011	(0.004) ***			0.012	(0.004) ***
ln*AGE*	0.298	(0.119) **	-0.110	(0.048) **	0.268	(0.110) **	-0.113	(0.048) **
MALE	0.535	(0.070) ***	-0.050	(0.030) *	0.529	(0.066) ***	-0.050	(0.030) *
PRIM	0.132	(0.104)	0.080	(0.042) *	0.135	(0.101)	0.085	(0.042) **
SECO	0.149	(0.142)	0.180	(0.048) ***	0.183	(0.130)	0.186	(0.048) ***
TERT	0.445	(0.172) ***	0.224	(0.050) ***	0.488	(0.158) ***	0.226	(0.049) ***
Q1	-0.549	(0.192) ***	-0.213	(0.068) ***	-0.570	(0.174) ***	-0.197	(0.066) ***
Q2	-0.445	(0.161) ***	-0.133	(0.064) **	-0.441	(0.149) ***	-0.114	(0.065) *
Q3	-0.511	(0.137) ***	-0.020	(0.065)	-0.490	(0.126) ***	-0.003	(0.064)
Q4	-0.413	(0.127) ***	0.003	(0.063)	-0.408	(0.120) ***	0.009	(0.062)
ln*HHS*	-0.034	(0.083)	-0.038	(0.041)	-0.030	(0.080)	-0.033	(0.040)
ln*MINCSB*	-0.015	(0.033)	0.008	(0.016)	-0.015	(0.029)	0.007	(0.016)
ln*ROOM*	0.129	(0.088)	0.065	(0.040) *	0.155	(0.078) **	0.070	(0.039) *
HOME	-0.034	(0.078)	0.032	(0.041)	-0.049	(0.072)	0.035	(0.041)
POWR	0.229	(0.140) *	-0.036	(0.069)	0.221	(0.133) *	-0.032	(0.068)
TAPW	0.239	(0.143) *	-0.055	(0.068)	0.222	(0.129) *	-0.055	(0.068)
TOIL	0.032	(0.093)	-0.016	(0.044)	0.023	(0.087)	-0.020	(0.044)
Constant	10.279	(0.601) ***	0.671	(0.247) ***	10.444	(0.522) ***	0.651	(0.236) ***
Obs.	858				858			
No. of district dummies	8		8		0		0	
Wald chi2	378.85				393.94			
Endogeneity test:								
C statistic chi2	8.579 ***				7.705 ***			
Overidentifying restriction test:								
Hansen's J stat.	4.48				5.70			

38. One may also be worried that the wage equations are different between the treated (i.e., commuters) and the untreated (i.e., noncommuters). The above treatment effects model implicitly assumes that they are the same. With the two groups differentiated, the treatment effects model can be rewritten by this:

$$\ln W_i = D_i(\alpha + X_i'\beta_1) + (1 - D_i)(X_i'\beta_2) + \varepsilon_i \qquad (1)'$$

D_i represents the treatment status and the selection mechanism is the same as Equation (2) above.

39. The results are shown in **Table 6**. There are slight differences in the wage equation between the treatment and control groups. For examine, the gender inequality in wages looks greater among noncommuters. That is, informal employees or family business workers are faced with a greater gender inequality, which may be plausible in the Malagasy economy. However, the results are broadly similar to the above. It can be concluded that the differences in the wage equations are negligible, and our main results shown above remain robust.

Table 6. Endogenous treatment effects model with different outcome equations between the treated and the untreated

Dependent var.	ln*W* for treated		ln*W* for non-treated		*D*	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
D [1]	1.027	(0.424) **				
ln*TAXIBE*					-0.080	(0.038) **
ln*SUBUBUS*					0.028	(0.043)
ln*MINCBD*					0.105	(0.188)
ln*RENT*					0.035	(0.012) ***
ln*AGE*	0.327	(0.148) **	0.285	(0.147) *	-0.359	(0.158) **
MALE	0.288	(0.073) ***	0.642	(0.094) ***	-0.151	(0.092) *
PRIM	0.007	(0.134)	0.150	(0.125)	0.296	(0.171) *
SECO	0.043	(0.157)	0.170	(0.171)	0.613	(0.173) ***
TERT	0.416	(0.179) **	0.463	(0.219) **	0.705	(0.172) ***
Q1	-0.425	(0.225) *	-0.533	(0.253) **	-0.733	(0.216) ***
Q2	-0.504	(0.156) ***	-0.406	(0.212) *	-0.422	(0.186) **
Q3	-0.483	(0.119) ***	-0.521	(0.181) ***	-0.064	(0.176)
Q4	-0.372	(0.111) ***	-0.429	(0.164) ***	0.006	(0.168)
ln*HHS*	-0.089	(0.087)	0.006	(0.104)	-0.100	(0.130)
ln*MINCSB*	-0.014	(0.038)	-0.003	(0.039)	0.024	(0.049)
ln*ROOM*	0.086	(0.089)	0.155	(0.108)	0.193	(0.117) *
HOME	-0.091	(0.085)	-0.028	(0.093)	0.093	(0.133)
POWR	0.201	(0.118) *	0.284	(0.184)	-0.109	(0.192)
TAPW	0.059	(0.131)	0.308	(0.172) *	-0.145	(0.188)
TOIL	0.079	(0.095)	0.048	(0.112)	-0.053	(0.132)
Constant	11.444	(0.546) ***	10.100	(0.766) ***	0.558	(0.783)
Obs.	858					
No. of district dummies	8		8		8	
Endogenous selection test:						

Wald chi2	5.24 *

[1] Average treatment effect is shown comparing the outcomes when D= 0 and D= 1.

40. Finally, the possible selection bias is examined. As discussed above, a number of household members were excluded from the above estimation. First, self-employees or those who engaged with family business were not asked to about their salaries during the survey. Second, people may not have reported their wages even if they were employed in the formal sector. There may be a systematic difference between those who answered their wages and those who did not. This raises a concern about the sample selection bias.

41. To deal with this problem, the conventional selection bias model is examined (e.g., Heckman, 1979; Lee, 1978). The following selection mechanism is considered:

$$d1 = \begin{cases} 1 & \text{if } Z_i'\gamma\theta_2 + X_i'\delta\theta_2 > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

where $d1$ is set to unity if a household member is employed in the formal sector and reported her wage level in the survey. The same independent and instrumental variables are used as the above. To avoid further complication, the commuting choice variable, D, is replaced with its predicted value, $\widehat{D}$. That is, our baseline model is the two-stage IV regression, which is consistent since D is endogenous given the above discussion.

The estimation result indicates that the sample selection bias does not seem to matter to our data (**Table 7**). The selection bias test statistic is estimated at 1.75, which cannot reject the hypothesis that there is a sample selection bias in the data, although there are some characteristics shared by household members who chose to work in the formal sector and answered their wages in the survey. For instance, those who are young are less likely to be employed in the formal sector: The coefficient of *AGE* is -0.342. Those who completed the tertial level of education are more likely to select formal jobs: The coefficient of TERT is estimated at 0.253. Other than the first stage regression, the result from the wage equation is

consistent with the baseline model without sample selection, confirming that the above estimation results are robust.

Table 7. Heckman sample selection model

	IV regression		Heckman selection		First stage regression	
Dep. Var.	ln*W*		ln*W*		*d*1	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
ln*TAXIBE*					0.004	(0.023)
ln*SUBUBUS*					0.041	(0.024) *
ln*MINCBD*					0.008	(0.101)
ln*RENT*					-0.021	(0.007) ***
$\widehat{D}$	3.566	(1.360) ***	1.442	(0.485) ***		
ln*AGE*	0.336	(0.118) ***	0.233	(0.107) **	-0.342	(0.085) ***
MALE	0.548	(0.067) ***	0.542	(0.060) ***	0.060	(0.047)
PRIM	0.069	(0.113)	0.148	(0.097)	0.048	(0.081)
SECO	0.052	(0.157)	0.155	(0.121)	0.023	(0.086)
TERT	-0.069	(0.342)	0.471	(0.154) ***	0.253	(0.094) ***
Q1	-0.481	(0.194) **	-0.524	(0.162) ***	0.189	(0.113) *
Q2	-0.464	(0.130) ***	-0.418	(0.131) ***	0.064	(0.106)
Q3	-0.518	(0.105) ***	-0.485	(0.105) ***	0.162	(0.098) *
Q4	-0.486	(0.103) ***	-0.372	(0.094) ***	0.117	(0.092)
ln*HHS*	-0.052	(0.068)	-0.041	(0.067)	-0.038	(0.070)
ln*MINCSB*	0.001	(0.025)	-0.010	(0.024)	0.029	(0.031)
ln*ROOM*	0.184	(0.066) ***	0.126	(0.067) *	-0.167	(0.066) **
HOME	-0.039	(0.065)	-0.041	(0.064)	-0.116	(0.073)
POWR	-0.018	(0.125)	0.229	(0.102) **	-0.124	(0.116)
TAPW	0.279	(0.106) ***	0.224	(0.098) **	-0.128	(0.097)
TOIL	0.146	(0.078) *	0.064	(0.070)	0.094	(0.075)
Constant	10.216	(0.578) ***	10.307	(0.470) ***	0.289	(0.434)
Obs.	858		2883			
No. of district dummies	8		8		8	
R squared	0.4346					
F stat.	26.38					
Wald chi2			659.15			
Selection test: Wald chi2			1.750			

VII. CONCLUSION

42. In recent years, there has been renewed interest in access to jobs in relation to transport connectivity. In Sub-Saharan Africa, about 14 million working age people are added to the

labor market every year. Ensuring sustained access to jobs seems to be a prerequisite for inclusive and robust economic growth. In theory, the people's demand for commuting is generally determined by marginal transportation costs and where they live. And the residential locational choice is dependent on housing price gradients.

43. The current paper aims at examining the impact of public transit connectivity on access to jobs by comparing wages earned by those who commute to work and those who do not commute, using the data from Antananarivo, the capital of Madagascar, which is experiencing rapid urbanization. It is found that the wages earned by commuters were systematically higher than those earned by those who decided not to commute and presumably are self-employed or engaged with family businesses around their neighborhood. Thus, the accessibility to jobs is important to increase household income and stimulate economic growth in the city.

44. The proximity to public transport, especially taxi-be, is instrumental to supporting the people's commuting choice. Of course, other factors, such as human capital, i.e., experiences and education, also matter to earn more. In Antananarivo, importantly, the analysis shows that there is a substantial gender inequality in wages. Women are more likely to use buses to commute, and yet, they earn less than men. In addition, the poor tend to benefit less from public transportation. The public bus services are affordable, however, they may be unavailable around their residential areas or practically inefficient because of their slow operating speed and unsafety. There are many policy issues to be improved.

REFERENCES

Angel, Shlomo, Patrick Lamson-Hall, Manuel Madrid, Alejandro Blei, and Jason Parent. 2016. *Atlas of Urban Expansion: The 2016 Edition*.

Atack Jeremy, and Robert Margo. 1998. "Location, location, location!" The price gradient for vacant urban land: New York, 1830 to 1990. *Journal of Real Estate Finance and Economics*, Vol. 16(2), pp. 151-172.

Battese, George. 1997. "A Note on the Estimation of Cobb-Douglas Production Functions when some Explanatory Variables have Zero Values." *Journal of Agricultural Economics*, vol. 48(2), pp. 250–252.

Bento, Antonio, Maureen Cropper, Ahmed Mushfiq Mobarak, and Katja Vinha. 2005. The effects of urban spatial structure on travel demand in the United States. *Review of Economics and Statistics*, Vol. 87(3), pp. 446-478.

Boudier-Bensebaa, Fabienne. 2005. Agglomeration economies and location choice: Foreign direct investment in Hungary. *Economics of Transition*, Vol. 13(4), pp. 605-628.

Ceccato, Vania. 2017. Women's victimisation and safety in transit environments. *Crime Prevention and Community Safety*, Vol. 19(3-4), pp. 163-167.

Chen, Jianquan, and Luca Bertolini. 2013. Measuring urban job accessibility with distance decay, competition and diversity. *Journal of Transport Geography*, Vol. 30, ppo. 100-109.

Cieślik, Andrzej, and Michael Ryan. 2004. Explaining Japanese direct investment flows into an enlarged Europe: A comparison of gravity and economic potential approaches. *Journal of the Japanese and International Economies*, Vol. 8(1), pp. 12-37.

Gedal, Michael, and Ingrid Gould Ellen. 2018. Valuing urban land: Comparing the use of teardown and vacant land sales. *Regional Science and Urban Economics*, Vol. 70, pp. 190-203.

Girsberger, Esther, and Romuald Meango. 2017. The puzzle of educated unemployment in West Africa. Unpublished paper.

Gobillon, Laurent, Dominique Meurs, and Sebastien Roux. 2015. Estimating gender differences in access to jobs. *Journal of Labor Economics*, Vol. 33(2), pp. 317-363.

Heckman, James. 1979. Sample selection bias as a specification error. *Econometrica*, Vol. 47(10), pp. 153–161.

Hensher, David. 1986. Sequential and full information maximum-likelihood-estimation of a nested logit model, *Review of Economics and Statistics*, Vol.68(4), pp. 657-667.

Hensher, David, and William Greene. 2002. Specification and estimation of the nested logit model: Alternative normalisations. *Transportation Research: Part B: Methodological* Vol. 36(1), pp. 1-17.

Hymel, Kent. 2009. Does traffic congestion reduce employment growth? *Journal of Urban Economics*, Vol. 65, pp. 127-135.

Iimi, Atsushi. 2019. Location, location, location revisited: Evidence from Antananarivo, Madagascar. Forthcoming.

ILO. 2016. World Employment and Social Outlook: Trends 2019. International Labour Organization.

Jin, Jangik, and Peter Rafferty. 2017. Does congestion negatively affect income growth and employment growth? Empirical evidence from US metropolitan regions. *Transport Policy*, Vol. 55, pp. 1-8.

Lee, Lung-Fei. 1978. Unionism and wage rates: a simultaneous equation model with qualitative and limited dependent variables. *International Economics Review*, Vol. 19, pp. 415-433.

Lozano-Gracia, Nancy, and Young Cheryl. 2014. Housing consumption and urbanization. Policy Research Working Paper 7112. World Bank.

Matas, Anna, Jose-Luis Raymond, and Jose-Luis Roig. 2009. Car ownership and access to jobs in Spain, *Transportation Research Part A*, Vol. 43, pp. 607-617.

McFadden, Daniel. 1974a. The measurement of urban travel demand. *Journal of Public Economics*, Vol. 3(4), pp. 303-328.

McFadden, Daniel. 1974b. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, edt. by P. Zarembka, pp. 105-142. Academic Press.

Mills, Edwin. 1972. *Studies in the Structure of the Urban Economy*. Johns Hopkins University Press.

Milner, Chris, Geoff Reed, and Pawin Talerngsri. 2006. Vertical linkages and agglomeration effects in Japanese FDI in Thailand. *Journal of The Japanese and International Economies*, Vol. 20(2), pp. 193-208.

Mondschein, Andrew, and Brian Taylor. 2017. Is traffic congestion overrated? Examining the highly variable effects of congestion on travel and accessibility. *Journal of Transport Geography*, Vol. 64, pp. 65-76.

Muth, Richard. 1969. *Cities and Housing*. University of Chicago Press.

Saghapour, Teyebeh, Sara Moridpour, and Russell Thompson. 2016. Public transport accessibility in metropolitan areas: A new approach incorporating population density. *Journal of Transport Geography*, Vol. 54, pp. 273-285.

Salon, Deborah. 2009. Neighborhoods, cars, and commuting in New York City: A discrete choice approach. *Transportation Research Part A*, Vol. 43, pp. 180-196.

UN-Habitat. 2016. *World City Report 2016: Urbanization and Development: Emerging Futures.*

Wang, Chih-Hao, and Na Chen. 2015. A GIS-based spatial statistical approach to modeling job accessibility by transportation mode: Case study of Columbus, Ohio. *Journal of Transport Geography*, Vol. 45, pp. 1-11.

World Bank. 2017. *Greater Antananarivo: Urban poverty and resilience study.* Washington DC: The World Bank.