Firm Formation and Occupational Changes in Tanzania

Daniel Lara-Agudelo and Varsha Sabhnani

April 26, 2019

1 Introduction

Tanzania has experienced rapid growth since the 1990s, with a GDP growth rate of about 5-7% since 2000, making it one of the highest performing countries in sub-Saharan Africa. However, this high growth rate has not translated into a proportional decline in poverty rates[1] (Page, 2016). During the same period, Tanzania has been transitioning from an agricultural economy to a service economy with very little growth in the manufacturing sector. Despite its small size, the manufacturing sector is the most productive with the potential to drive the future growth of the country. As a result, the Tanzanian government has set initiatives to encourage the development of industry. Given that poverty has been decreasing slowly in comparison to economic growth, it is important to ask whether, and to what extent, the growth of the industrial sector will a ect Tanzanians.

In this paper we take the rst step towards answering this question by investigating the e ects of the opening of manufacturing rms[2] on the employment decisions of Tanzanians. Much has been written about the relationship between rm formation and regional development, however, the bulk of this literature is restricted to research on the developed world. This is probably because the data required for such studies can only be found in countries with detailed records. Our novel contribution will be to modify the framework described in the literature to measure the responses to rm formation in a developing country.

Workers in developing countries are more responsive to new employment opportunities. Thus, we can expect that the opening of a rm in this setting will have di erent impacts than the opening of one in a developed country where people's jobs are more stable and the country has already been industrialized. As employment in industry tends to pay more than employment in agriculture, the opening of a rm would raise the opportunity cost of remaining in this sector. Therefore, we should expect to see workers moving out of agriculture and into industry. Additionally, the higher average wage will increase local demand, which

[1] It is worth noting that GDP growth is overstated unless we take population growth into account. Tanzania is about average in terms of GDP per capita when compared to other sub-Saharan countries.

[2] The rest of this section uses manufacturing rms and rms interchangeably

will create an incentive to start new rms to provide additional goods and services. We would then expect workers to move out of agriculture and into services, as the service sector is more productive.

Gender is another important dimension to consider in this setting. The mining literature in Tanzania shows that the opening of a new mine attracts men into mining, and women into services (Chuhan-Pole, Dabalen & Land, 2017). As the roles of men and women are more rigidly di erentiated in developing countries, we can expect men and women to react di erently to a rm opening.

Additionally, in developing countries like Tanzania, the decision making unit is often the household as opposed to the individual, and households tend to use diversi cation to cope with poverty (Ellis, 2000). This is to say that in a given household some members might be engaged in agriculture while others are engaged in di erent sectors to mitigate the negative e ects of a shock in any given sector. The opening of a new rm can provide a new avenue for diversi cation. However, if it is the case that the additional wages from industry substantially raise the household'sTd [(sda)27(v)272d [(sd72d (of)e)-444(T)837ba-27(e)-383(work

in agriculture for both men and women, and increases employment in both services and industry. The e ect of rm formation on overall employment is found to be ambiguous.

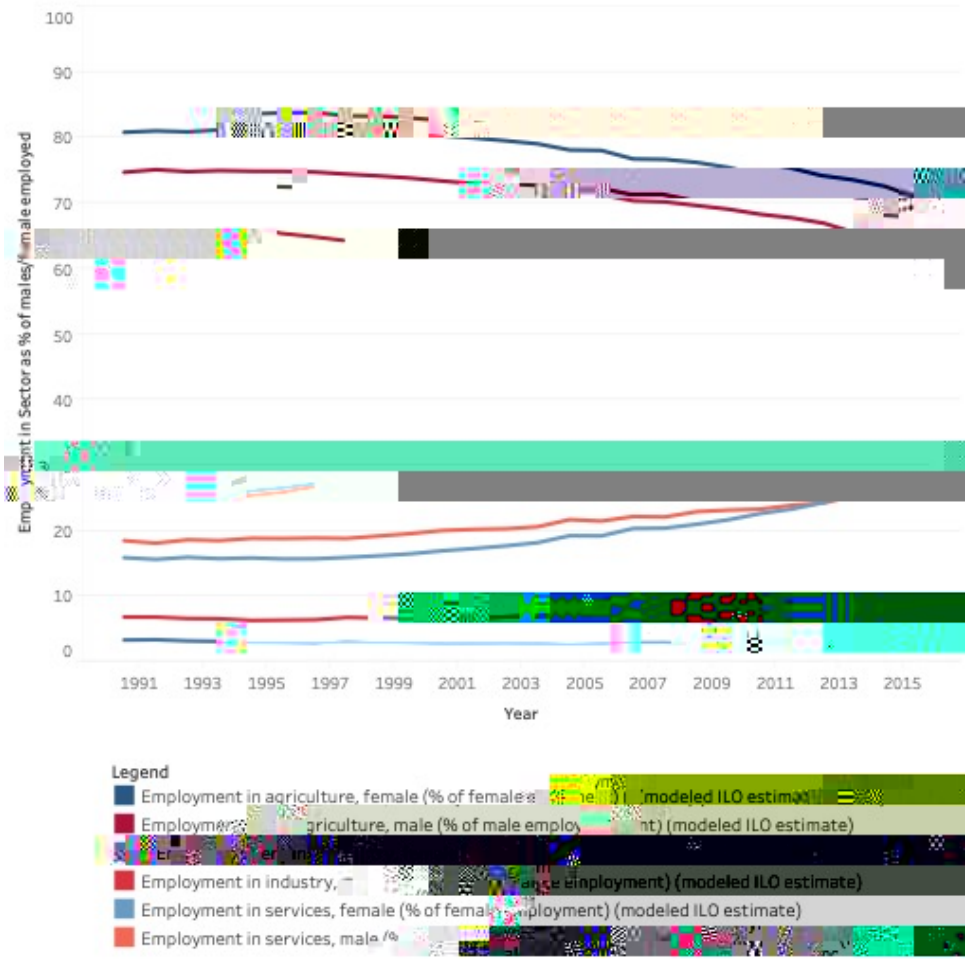

Figure 1: Employment of men and women by sector over time.

2.2 Size and Composition of the Manufacturing Sector

While the manufacturing sector in Tanzania only employs around 6% of Tanzanians, it is important to note that it contributed to 28.9% of GDP in 2017. In comparison, agriculture contributed to 23.4% of GDP in the same year, while services contributed to 47.6% of GDP (CIA, 2019). It is clear that the industrial sector is more productive than the other two, and contributes disproportionately to GDP given its size. It is mostly dominated by agro-processing (55%), followed by furniture (13%), non-metallic mineral products (11%), tobacco (7%), and textiles (5%) (Page, 2016). Tanzania has an extremely high rate of entrepreneurship (25%) and approximately 18% of Tanzania's household enterprises are

involved in manufacturing, especially in the manufacturing of beverages, apparel and furniture. (Page, 2016).

3 Literature Review

3.1 Evidence from Mining in Tanzania

There are no papers that directly investigate the regional effects of manufacturing firm formation in Tanzania. However, there are papers investigating the socio-economic effects of mine openings. While the dynamics surrounding mine openings and industrial firm formation are different, examining the mining literature can give us a general sense of employment trends in Tanzania.

Even though a mine may not directly increase employment as it is very capital intensive, its opening or restarting generates a clustering of economic activity around it. Wages in mining tend to be higher than the average wages in the community, which results in a higher demand for goods and services from those who work in mines. This would hypothetically result in the reallocation of labor and other resources into new sectors like services and away from agriculture. Additionally, it is often hypothesized that access to resources like water, electricity, or toilets might increase as a result of mine openings (Chuhan-Pole et al., 2017).

The authors of Mining in Africa (Chuhan-Pole et al., 2017) find that women have much higher non-farm opportunities, especially in sales and services, near an active mine site. Women's employment in agriculture is found to decline, while their likelihood to work year long rises. Overall employment is found to rise but employment in agriculture is found to decrease. They also find evidence to support the idea that opening a mine causes indirect positive effects, such as increased regional economic growth and increased access to electricity. They note, however, that the increased regional growth is not persistent and the effect dissipates over time.

3.2 Firm Formation on Regional Employment

employment. Thus, the effect of a firm opening will change over time, and its overall impact will depend on the relative sizes of the positive and negative pressures. It is important to note here that this mechanism can be observed even if the new firm does not succeed, and that the magnitude of its effect will depend on its quality, i.e. how efficient it is. This framework also assumes a well-formed and reactive market which forces existing producers to become more efficient in the face of a new producer.

The caveats to this model are particularly important to our study as the manufacturing sector in Tanzania is relatively small and not very competitive (Page, 2016). This means that we should expect to see minimal changes in employment due to supply-side effects, and mostly observe the direct employment effects.

In his empirical work, Fritsch uses an adjusted start-up rate as the independent variable, and a 2-year change in employment as the dependent variable. For this paper, we do not have access to firm start-up rates. Instead, we use the starting year of firms to create an index to capture the size and number of firms in the district.

regional effects. An important finding from this paper is that the overall effect in the first six to seven years may be negative.

This result is confirmed by Mueller, van Stel and Storey (2007) in their paper on firm formation and its effects of regional development in Great Britain. In particular, they look at the difference in effects of firm formation in low-enterprise and high-enterprise counties and find that the effect of firm formation on employment in low-enterprise counties is negative.

Delfmann and Sierjdan (2014) use panel data from the Netherlands to look at the effects of firm formation on employment growth in declining regions using the Fritsch model. We are particularly interested in their findings for rural areas as about two-thirds of Tanzanians live in a rural setting. Comparing the effects of firm formation in rural and urban regions, they find that there are large effects of new firms in rural areas, and that most of these effects are due to the large positive initial effects. We can then expect that the main driver of employment change will be due to the immediate effects of firm formation.

4 Data

This study employs two data sets, each created by combining data on manufacturing firms with household data either from the demographic and health surveys or the decennial census. Firm formation in Tanzania was captured using the Census of Industrial Production (CIP) which was published by the National Bureau of Statistics (2013). The CIP provides an exhaustive list of all existing firms in mainland Tanzania with more than ten employees[3].

these 1,084 manufacturing rms which will form the basis of our analysis[4]. It is evident that rms are not randomly distributed within the country. We attempt to mitigate the resulting bias by running separate regressions by rural-urban status and omitting data points in major cities as rms tend to locate in urban areas. However, given the data, it is impossible to control for this e ect altogether.

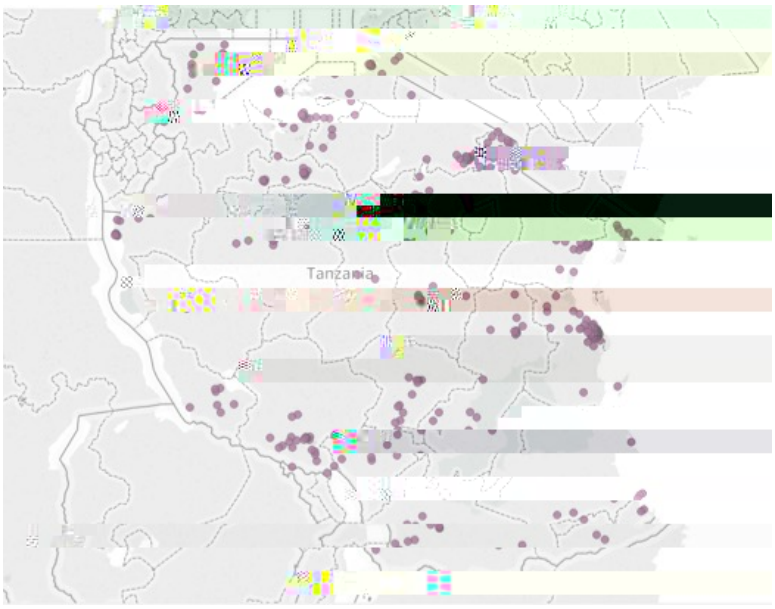

Figure 2: Distribution of manufacturing rms employing more than 10 workers.

The independent variable in both analyses was constructed using the data from the CIP. Our goal was to create an index which captured the level of rm formation in a given area. We wanted to account not only for the number of rms that were opened but also their size. Unfortunately, the census did not report the exact number of workers employed at a given rm. Instead it categorized them into ve size classes: 10-19, 20-49, 50-99, 100-499, and 500+ employees. Additionally, we wanted to account for the number of rms already in existence in a district. It is undoubtedly the case that a single large rm opening in an area with several rms in operation has a lesser impact when compared to the same rm opening in an area with no preexisting rms. These considerations led us to create the index in the

[4] See Appendix for maps showing the geo-spatial distribution of rms by size.

following manner:

$$\text{index} = \frac{\sum_{i=1}^{5} s_i x_i}{\log \sum_{i=1}^{5} s_i y_i}$$

where s_i is the minimum bound of size class i and x_i and y_i are the number of new and prexisting firms belonging to that size class in a given district, respectively. This was modeled after the index used in Fritsch (1997) with the addition of the natural log in the denominator and an adjustment for the different sizes of firms. We found that without this adjustment the index would overemphasize the dampening effect of having a certain number of firms already operating in the area.

To get a more complete picture of the impact that industrialization has on the employment options of Tanzanians, we conducted two distinct analyses using data from different surveys. The first was at the individual level and used a series of demographic and health surveys (DHS) created by the National Bureau of Statistics and funded by USAID. The DHS program collects information on the health and welfare of women and children in the developing world. In Tanzania, thirteen of these surveys have been implemented since 1991, most of which are used to track the spread of Malaria and AIDS. However, they contain enough information for us to be able to analyze the occupational choices of men and women living in the country. It should be noted that women are slightly over-sampled due to the goals of the survey. Since the analysis required that we match respondents and firms according to their location, we could only use surveys that included a geographic identifier for everyone in the sample[5]. Thus, our data set consists of individuals who participated in the following DHS waves: 1999, 2003/2004, 2007/2008, 2009/2010 and 2011/2012[6]. The only conditions imposed on the sample were that respondents be in their prime working age and living in the mainland[7]. Furthermore, in order to preserve the independence of the observations, we selected only

[5] The government made several changes to the country's administrative boundaries in both 2002 and 2012. In order to accurately match firms and respondents to the correct district and region we geocoded the observations from the CIP and DHS and passed the resulting coordinates through a shapefile containing the most recent borders.

[6] Some of the surveys were implemented over the course of two years.

[7] We consider an individual to be of prime working age if they are older than 18 and younger than 54.

Table 1: Summary statistics for the female and male sample of the demographic and health surveys.

Variables	1999		2003/2004		2007/2008		2009/2010		2011/2012		Total	
Controls	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Age	32.25	36.25	31.93	34.28	32.79	35.33	33.09	35.44	33.40	35.80	32.81	35.31
Years of Schooling	4.48	5.81	4.93	6.23	5.08	6.18	5.23	6.23	5.37	6.48	5.11	6.25
Household Size	5.76	5.16	5.61	5.04	5.73	5.31	5.66	5.01	5.96	5.39	5.75	5.22
Rural	0.68	0.68	0.75	0.76	0.79	0.80	0.77	0.78	0.79	0.78	0.77	0.77
Outcomes												
Employed	0.87	0.97	0.92	0.99	0.91	0.99	0.88	0.98	0.93	0.99	0.91	0.99
Agriculture												
of Total	0.62	0.64	0.71	0.67	0.69	0.71	0.65	0.66	0.73	0.72	0.69	0.69
of Employed	0.72	0.66	0.77	0.68	0.76	0.72	0.73	0.67	0.79	0.72	0.76	0.70
Services												
of Total	0.03	0.03	0.16	0.15	0.17	0.10	0.05	0.06	0.07	0.06	0.10	0.09
of Employed	0.04	0.03	0.18	0.15	0.19	0.10	0.05	0.06	0.08	0.06	0.11	0.09
Professional												
of Total	0.02	0.05	0.03	0.04	0.03	0.07	0.02	0.05	0.02	0.07	0.03	0.06
of Employed	0.03	0.06	0.03	0.04	0.03	0.07	0.02	0.05	0.03	0.07	0.03	0.06
Manufacturing/Trade												
of Total	0.19	0.25	0.02	0.13	0.02	0.12	0.17	0.22	0.10	0.15	0.09	0.16
of Employed	0.22	0.25	0.02	0.13	0.02	0.12	0.19	0.22	0.11	0.15	0.10	0.16
Observations												
Total	1,685	1,301	4,167	2,930	3,961	2,497	4,678	964	5,351	3,377	19,842	11,069

model but instead of a binary treatment variable we used the index to measure the intensity of treatment. The treatment group was made up of districts where new rms opened at some point between the two censuses. The control group consisted of districts where no new rms were formed. It was important to account for intensity because it was seldom the case that a district saw only a single rm start up. The model is therefore given as follows:

$$Y_{it} = {}_0 + {}_1 Index_i \quad After_t + {}_2 Index_i + {}_3 After_t + {}_4 X_{it} + {}_i + {}_{it} \qquad (2)$$

where *After* is a binary variables that describes whether a particular observation was taken from the 2012 census (after treatment) or taken from the 2002 census (before treatment). As before, X_{it} is a vector of controls which includes the average age and years of schooling for respondents in a given district as well as the proportion of people living in a rural area.

5.3 Limitations

Of course, no econometric model is perfect and given the quality of the data that is available on developing countries like Tanzania, there are various limitations to our approach. One such limitation that was touched upon in the previous section has to do with the fact that

tend to be quite small. Tanzania in particular has a high rate of self-employment due to the lack of jobs, but the rms created as a result of this self-employment are generally small and unproductive. Since we are examining only rms with more than 10 employees, this should not introduce too much bias into our results. Lastly, data on migration was unavailable in both the census and the DHS which made it di cult to control for the fact that people may be moving as new rms start to open up in other districts. Tanzanians are very mobile; it is normal for young men to go to the city to nd a job that supports their family back home. Not controlling for this could possibly bias our results. However, since we sampled only heads of households and their spouses it is unlikely that omitted variable bias is particularly severe in this case since it is usually older male sons who move to nd work while the heads stay home to look over the family's agricultural plot.

6 Results and Discussion

Both of our analyses indicated that men and women react similarly when new rms commence operations in their districts. Regardless of gender, workers left the agricultural sector and transitioned into the more productive services and manufacturing sectors. Concerning the magnitude of the e ect,ring 0 iy-1125(Res0 iyw [(sev)270 iys(w)w8(t)1(the)ix(samp9(h)7(uutamp9(h

Table 3: Regression results for the DHS models of employment outcomes on index for women.

	Employed	Agriculture	Services	Professional	Manufacturing/ Trade
Index	-0.0131***	-0.0530***	0.0226***	-0.0170**	0.0180***
	(0.00197)	(0.00539)	(0.00243)	(0.00671)	(0.00253)
Age	0.0990***	-0.143***	0.175***	0.0967	0.0401
	(0.0305)	(0.0271)	(0.0337)	(0.0812)	(0.0336)
Age^2	-0.000667	0.00210***	-0.00253***	-0.000892	-0.000836*
	(0.000471)	(0.000402)	(0.000494)	(0.00119)	(0.000501)
Years of Schooling	0.0535***	-0.188***	0.0601***	0.508***	0.0233**
	(0.00922)	(0.00930)	(0.00987)	(0.0226)	(0.00942)
Household Size	0.0150	0.106***	-0.0808***	-0.0141	-0.0743***
	(0.0101)	(0.0113)	(0.0132)	(0.0279)	(0.0133)
Rural	1.257***	2.523***	-2.079***	-0.786***	-1.998***
	(0.0697)	(0.0598)	(0.0733)	(0.147)	(0.0760)
2003/2004	0.733***	0.195*	2.423***	0.0776	-2.725***
	(0.121)	(0.107)	(0.221)	(0.261)	(0.157)
2007/2008	0.512***	-0.0697	2.669***	0.0623	-2.775***
	(0.122)	(0.107)	(0.222)	(0.269)	(0.166)
2009/2010	0.343***	-0.143	0.799***	-0.591**	0.0410
	(0.118)	(0.103)	(0.230)	(0.259)	(0.0986)
2011/2012	0.636***	0.224**	1.340***	-0.456*	-0.618***
	(0.122)	(0.106)	(0.229)	(0.273)	(0.106)
Constant	-1.763***	2.194***	-5.406***	-8.716***	-0.237
	(0.476)	(0.431)	(0.583)	(1.350)	(0.531)
Observations	19,842	18,013	18,013	18,013	18,013
R^2	0.098	0.353	0.264	0.356	0.268

This suggests that the entry of manufacturing firms, whose workers are more productive and earn higher wages, raises the opportunity cost of women staying in the primary sector. Previous research shows that it is not difficult to enter the services and manufacturing sectors, so it is not surprising to see that women respond in this way to these new wage incentives. Looking again at Table 3 we note that firm formation has a larger impact on employment in services as compared to manufacturing, even though the relationship is positive for both. This confirms findings in the mining literature where researchers have found that the influx of workers into a new mining town increases the demand for certain services that women are more apt to provide. The same effect might explain what we see here with firm formation. Lastly, note that the change in the employment of women is negative in the regression table. We believe that this may be indicative of the fact that, because womens' employment decisions are made in conjunction with those of the other household members, higher wages in manufacturing allow males to earn more and give women the opportunity to stay home. Evidence that this is indeed the case comes from observing the coefficient on the variable household size which tells us that women with larger families are less likely to be employed in services and manufacturing and more likely to work in agriculture (often at home).

Looking now at the results from the male sub-sample of the DHS (Table 4), we find that the effects of firm formation on employment outcomes are largely the same as what we saw for females. While the likelihood of being employed in general and in agriculture are negatively related to increases in firm formation, the opposite is true when it comes to services and manufacturing. One interesting difference between the sexes is that men were found to be relatively more likely to work in manufacturing compared to services, while the opposite was true for women. In the mining literature researchers found similar results whereby the opening of a mine in a particular area increased male employment in mining and female employment in services. One unexpected finding from the regression output was that firm formation decreased the likelihood of men being employed overall. There are various reasons for why this may be the case including the fact that new firms increase the level of

Table 4: Regression results for the DHS models of employment outcomes on index for men.

	Employed	Agriculture	Services	Professional	Manufacturing/ Trade
Index	-0.0109**	-0.0722***	0.0129***	0.00556	0.0148***
	(0.00499)	(0.00920)	(0.00262)	(0.00417)	(0.00235)
Age	0.378***	0.0377	-0.0205	-0.0406	0.0229
	(0.109)	(0.0376)	(0.0467)	(0.0713)	(0.0399)
Age^2	-0.00515***	-0.000223	3.02e-05	0.000945	-0.000678
	(0.00151)	(0.000529)	(0.000663)	(0.000986)	(0.000561)
Years of Schooling	-0.0770*	-0.231***	0.0309***	0.314***	0.0382***
	(0.0396)	(0.0121)	(0.0117)	(0.0220)	(0.0103)
Household Size	0.0522	0.0543***	-0.0214	-0.0483**	-0.0416***
	(0.03913504(b360((0.0399))]ie-D46d0500.046o01())-2305((0.0)1(2))-2Rur776-7601892				

competition in the local economy making it more di cult for small rms, and therefore their owners, to survive in the market. Unfortunately, we cannot test this hypothesis without more detailed data on entrepreneurship in Tanzania although qualitative descriptions of the business climate in the country suggest that it is common for people to start their own rms.

We see the same patterns as before when looking at results from the census data analysis with a few notable exceptions. It is important to note while reading this portion of the paper that while the two models capture the same general e ect they are not strictly comparable. The rst analysis looks at the likelihood that an individual chooses a particular occupation while the second examines the proportion of people in a district employed in a given sector. Another point of importance is that the DHS models look at rms opening in the previous year while the census model measure the e ect of the treatment (i.e. the number of rms) opening within a ten-year period.

Table 5: Census regression results of employment outcomes on index for women.

	Employed	Labor Force Participation	Agriculture	Services	Professional	Manufacturing/ Trade
After	-0.00296	-0.00544	-0.0949***	0.0153	0.0363***	0.0114**
	(0.00462)	(0.0328)	(0.0206)	(0.0115)	(0.00827)	(0.00498)
After * Index	0.000195***	0.000371***	-0.000198**	0.000136*	-3.10e-05	0.000156***
	(2.33e-05)	(0.000113)	(8.66e-05)	(7.27e-05)	(5.02e-05)	(4.57e-05)
Age	0.00626**	0.0203	0.0636***	-0.0237***	-0.0182***	-0.00517
	(0.00308)	(0.0147)	(0.0120)	(0.00604)	(0.00661)	(0.00354)
Years of Schooling	-0.00321	-0.0325	-0.0213	0.0124	0.00599	0.00612
	(0.00494)	(0.0271)	(0.0159)	(0.0129)	(0.00697)	(0.00553)
Rural	0.0325	0.164**	0.299***	-0.199***	-0.0465	-0.0281
	(0.0214)	(0.0775)	(0.0704)	(0.0468)	(0.0292)	(0.0213)
Constant	0.773***	0.181	-1.288***	0.920***	0.631***	0.176
	(0.111)	(0.526)	(0.419)	(0.241)	(0.219)	(0.136)
C600.526)ening	-squly					

As in the DHS regressions, the census regressions for women (Table 5) report a negative relationship between levels of firm formation and agricultural employment. In addition, we see that female employment in services and manufacturing increases as new firms open in a given district. Interestingly, the results from the second model depart from those of the first in that the size of the employment change for women in manufacturing is larger than that for services. While it is not immediately obvious why we are seeing this change, it could be because of the addition of district-fixed effects in the second model. This would suggest that the level of female employment in the services and manufacturing sector is dependent on location. Another major difference between the two models is the increase in overall employment for women associated with an increase in the number of firms. Since the census is a larger data set and it captures the employment rate more accurately, we give more weight to these results. One of the benefits of using the census data is that we can measure labor force participation. For women in particular we see a positive effect of firm formation on labor force participation. This may be an indication that the opportunity cost of not working increases when new jobs are available in more productive sectors like manufacturing or services.

In the census regressions for males (Table 6) we see an increase in overall employment, as well as an increase in employment in the manufacturing sector as a result of new firm formation. However, in contrast to the DHS results, employment in services has a negative coefficient. This goes against the predictions we made earlier in the paper where we expressed the belief that both manufacturing and services employment would increase regardless of sex. Interestingly, the coefficient on agriculture is insignificant even though there is a clear decrease in the proportion of men who are employed in agriculture between 2002 and 2012 (Table 2). As we mentioned at the start of this section, an important difference between the census and the DHS models is that the latter captures the effects of firm formation after a year and the former captures the effect over 10 years. In other words, we are looking at averaged out effects in the second model (census), but more direct effects in the first model

Table 6: Census regression results of employment outcomes on index for men.

	Employed	Labor Force Participation	Agriculture	Services	Professional	Manufacturing / Trade
After	0.00817	-0.0403***	-0.0485**	-0.00768	0.0239***	0.00616
	(0.00813)	(0.0121)	(0.0194)		(0.00024)	

Table 7: Regression results for both models by urban-rural status

	Rural					
	Employed	Labor Force	Agriculture	Services	Professional	Manufacturing/ Trade
Females (DHS)	-0.0390***	-	-0.0383***	0.0338***	-0.0574	0.0352***
	(0.00481)		(0.00624)	(0.00615)	(0.0517)	(0.00651)
Males (DHS)	-0.0326**	-	-0.0539***	0.0383***	0.0329***	0.0200***
	(0.0135)		(0.0101)	(0.00603)	(0.00781)	(0.00514)
Females (Census)	-4.85e-06	-0.000255	-6.77e-05	0.000338	2.37e-05	0.000110*
	(5.46e-05)	(0.000408)	(0.000291)	(0.000306)	(0.000141)	(5.75e-05)
Males (Census)	-2.75e-05	-0.000398	-9.75e-06	0.000182	-3.06e-05	0.000373*
	(0.000104)	(0.000415)	(0.000158)	(0.000155)	(0.000141)	(0.000194)
	Urban					
	Employed	Labor Force	Agriculture	Services	Professional	Manufacturing/ Trade
Female (DHS)	-0.00877***	-	-0.0698***	0.0213***	-0.0120**	0.0186***
	(0.00206)		(0.00954)	(0.00267)	(0.00581)	(0.00273)
Male (DHS)	-0.00882*	-	-0.0986***	0.00882***	-0.00135	0.0151***
	(0.00466)		(0.0168)	(0.00271)	(0.00458)	(0.00256)
Female (Census)	0.000157***	0.000274*	1.84e-05	7.66e-05	-7.15e-05	1.24e-05
	(3.68e-05)	(0.000146)	(0.000128)	(0.000127)	(5.47e-05)	(3.31e-05)
Male (Census)	0.000126*	0.000143**	-0.000137	-2.32e-05	-2.87e-05	0.000169*
	(6.14e-05)	(5.91e-05)	(0.000171)	(9.75e-05)	(8.85e-05)	(8.64e-05)

Running the regressions separately for urban and rural areas with the census data did not yield many signi cant results. This is likely due to the fact that the data set was collapsed to district-level data, and very few districts are strictly rural or urban. For the purposes of this regression, we de ned a district as rural if more than 50% of the inhabitants of the district reported living in rural areas in both 2002 and 2012. The only signi cant coe cient for the rural districts under the census model were for manufacturing, which was positive and signi cant across both genders. These results are consistent with the original model. For

urban districts, the coe cients on overall employment and labor force participation were both positive and signi cant regardless of gender. These results are consistent with the original analysis, the only exception being labor force participation, which was originally insigni cant for women.

On the other hand, the urban-rural analysis for the DHS data yielded more signi cant results which were consistent with earlier ndings from the model. We see that rm formation has a negative e ect on overall employment as well as on employment in agriculture regardless of urban-rural status and gender. Similarly, rm formation has a positive e ect on employment in manufacturing across urban-rural status and gender.

Finally, regressions for both the models that excluded major urban areas such as Dar es Salaam were consistent with the results from the original regressions that we reported in the previous section across males and females.

6.1.2 Lags

A critique of some of the rm formation literature is that lags between rm formation and their e ects are not taken into account. However, the lack of competition in the manufacturing sector in Tanzania makes it likely that these lags are either small or do not exist since we do not expect to see the supply-side e ects predicted by Fritsch (1997). However, we wanted to check for varying e ects over time in case these supply-side e ects are created following rm formation.

The structure of the census data does not lend itself to the kind of lag analysis reported in the literature, but we were able to check for lagged e ects using the DHS data. We tested up to nine years of lags as suggested by the literature. We rst tested each lag separately, and then all together as in Fritsch and Mueller (2004).

We found that considering the e ect of rm formation *i* years ago, where *i* runs from 14).W9 moantoan

Hhe e ect ef rm eormation

6.1.3 Index Construction

We constructed the index by modifying the index in Fritsch (1997) to suit the needs of

in that district tend to move out of agriculture and into manufacturing and services regardless of sex. The e ect on overall employment was ambiguous. While the DHS model suggests that individuals are less likely to work following the formation of new manufacturing rms, the census model suggests the opposite. We also found con icting results with regards to employment in services for women. While the DHS results suggest that women tend to

8 Bibliography

Acs, Z.; Armington, C. (2003): Endogenous Growth and Entrpreneurship Activity in Cities, Washington DC: Center for Economic Studies, U.S. Bureau of the Census, CES 03-02.

9 Appendix

9.2 Regressions Using Log Index

Table 8: Regression results for the female-DHS analysis using log of index.

	Employed	Agriculture	Services	Professional	Manufacturing/ Trade
Index	-0.269***	-0.599***	0.373***	-0.0912	0.258***
	(0.0323)	(0.0338)	(0.0307)	(0.0604)	(0.0381)
Age	0.147***	-0.0776**	0.176***	-0.0130	-0.0466
	(0.0366)	(0.0302)	(0.0356)	(0.0713)	(0.0382)
Age^2	-0.00136**	0.00111**	-0.00252***	0.000830	0.000249
	(0.000564)	(0.000440)	(0.000521)	(0.00103)	(0.000558)
Years of Schooling	0.0398***	-0.187***	0.0355***	0.462***	0.0303**
	(0.0118)	(0.0106)	(0.0105)	(0.0224)	(0.0118)
Household Size	-0.00378	0.0829***	-0.0729***	-0.00911	-0.0448***
	(0.0113)	(0.0122)	(0.0140)	(0.0262)	(0.0156)
Rural	1.294***	2.330***	-1.994***	-0.748***	-1.702***
	(0.0959)	(0.0660)	(0.0820)	(0.152)	(0.0990)
2003/2004	-0.896***	0.257**	2.873***	-0.145	-2.847***
	(0.259)	(0.111)	(0.229)	(0.204)	(0.156)
2007/2008	-1.122***	0.0475	3.094***	-0.177	-2.914***
	(0.259)	(0.111)	(0.230)	(0.205)	(0.165)
2009/2010	2.648***	-0.412***	1.135***	-0.381	0.160
	(0.793)	(0.135)	(0.281)	(0.263)	(0.134)
2011/2012	-1.000***	0.308***	1.756***	-0.644***	-0.800***
	(0.259)	(0.110)	(0.234)	(0.220)	(0.109)
Constant	-0.608	1.447***	-5.976***	-6.509***	1.114*
	(0.610)	(0.499)	(0.614)	(1.199)	(0.633)
Observations	15,744	14,641	14,641	14,641	14,641
R^2	0.1380	0.3526	0.2735	0.3173	0.2673

Table 9: Regression results for the male-DHS analysis using log of index.

	Employed	Agriculture	Services	Professional	Manufacturing/ Trade
Index	-0.150***	-0.608***	0.214***	0.0101	0.233***
	(0.0441)	(0.0336)	(0.0323)	(0.0641)	(0.0276)
Age	0.134***	-0.0870***	0.0155	0.0500	0.0890***
	(0.0469)	(0.0306)	(0.0429)	(0.0741)	(0.0321)
Age^2	-0.00132*	0.00146***	-0.000459	-0.000436	-0.00149***
	(0.000716)	(0.000441)	(0.000614)	(0.00105)	(0.000466)
Years of Schooling	0.0602***	-0.215***	0.0575***	0.350***	0.0240***
	(0.0140)	(0.0101)	(0.0108)	(0.0231)	(0.00864)
Rural	0.845***	2.450***	-1.691***	-0.388***	-1.909***
	(0.111)	(0.0646)	(0.0920)	(0.142)	(0.0697)
Household Size	0.0321*	0.0853***	-0.0350**	-0.0521**	-0.0643***
	(0.0171)	(0.0121)	(0.0165)	(0.0246)	(0.0123)
2003/2004	2.379***	-0.0683	1.804***	0.0394	-0.805***
	(0.201)	(0.111)	(0.212)	(0.244)	(0.114)
2007/2008	3.263***	-0.0460	1.457***	0.551**	-0.759***
	(0.307)	(0.116)	(0.219)	(0.244)	(0.117)
2009/2010	0.348***	-0.122	0.721***	-0.293	0.0159
	(0.112)	(0.106)	(0.216)	(0.238)	(0.0976)
2011/2012	3.134***	0.121	0.610***	0.477*	-0.381***
	(0.283)	(0.114)	(0.227)	(0.255)	(0.110)
Constant	-2.036***	1.464***	-2.819***	-6.550***	-1.234**
	(0.735)	(0.503)	(0.730)	(1.255)	(0.522)
Observations	15,167	14,328	14,328	14,328	14,328
R^2	0.1997	0.3729	0.1940	0.2429	0.2019

Table 10: Regression results for the female-census analysis using log of index.

	Employment	Labor Force Participation	Agriculture	Service	Professional	Manufacturing/ Trade
After	-0.00558	-0.00960	-0.0904***	0.0137	0.0351***	0.00814
	(0.00466)	(0.0335)	(0.0208)	(0.0119)	(0.00835)	(0.00507)
After * Index	0.00449***	0.00748	-0.00694**	0.00291	0.00135	0.00503***
	(0.00108)	(0.00461)	(0.00350)	(0.00211)	(0.00171)	(0.000900)
Age	0.00706**	0.0223	0.0638***	-0.0230***	-0.0193***	-0.00516
	(0.00305)	(0.0143)	(0.0119)	(0.00598)	(0.00656)	(0.00341)
Years of Schooling	-0.00278	-0.0307	-0.0195	0.0129	0.00399	0.00512
	(0.00525)	(0.0270)	(0.0167)	(0.0130)	(0.00722)	(0.00550)
Rural	0.0414**	0.185**	0.299***	-0.192***	-0.0556*	-0.0263
	(0.0196)	(0.0761)	(0.0688)	(0.0462)	(0.0290)	(0.0204)
Constant	0.740***	0.0951	-1.304***	0.892***	0.680***	0.180
	(0.113)	(0.513)	(0.415)	(0.237)	(0.218)	(0.134)
Observations	206	206	206	206	206	206
R^2	0.344	0.365	0.859	0.725	0.696	0.728

Table 11: Regression results for the male-census analysis using log of index.

	Employment	Labor Force Participation	Agriculture	Services	Professional	Manufacturing/ Trade
After	0.00772	-0.0366***	-0.0481**	-0.00531	0.0218***	0.00408
	(0.00852)	(0.0123)	(0.0192)	(0.0105)	(0.00648)	(0.00820)
After * Index	0.00399**	-0.00510**	-0.000614	-0.00612***	0.00174	0.00743***
	(0.00158)	(0.00197)	(0.00384)	(0.00197)	(0.00181)	(0.00161)
Age	-0.000819	0.0194**	0.0410***	-0.0136**	-0.000654	-0.0143***
	(0.00539)	(0.00763)	(0.0119)	(0.00567)	(0.00528)	(0.00524)
Years of Schooling	-0.00133	-0.0453***	-0.0155	0.00931	0.00490	0.00654
	(0.0102)	(0.0109)	(0.0218)	(0.0144)	(0.00760)	(0.0101)
Rural	0.0482	-0.0403	0.381***	-0.179***	-0.0473	-0.128***
	(0.0331)	(0.0325)	(0.0801)	(0.0474)	(0.0338)	(0.0359)
Constant	0.955***	0.600**	-0.759*	0.608***	0.104	0.592***
	(0.210)	(0.261)	(0.426)	(0.227)	(0.190)	(0.213)
Observations	206	206	206	206	206	206
R^2	0.162	0.839	0.734	0.422	0.600	0.716