Using Micro Data to Measure Structural Change

Maggie McMillan
Tufts/IFPRI/NBER
Jan. 30, 2015

Cover Today

- DHS pros and cons
- An Example from Tanzania which incorporates all available sources of data

Introduction to Demographic and Health Surveys

- The DHS Program: Overview and survey types
- DHS survey topics
- DHS history and survey design
- DHS survey instruments
- DHS data sets
- Web tools
- What can DHS data tell about structural change ?

The DHS Program

- Program was established in 1984 by USAID
- The Monitoring and Evaluation to Assess and Use Results Demographic and Health Surveys (MEASURE DHS) project has provided technical assistance to more than 260 surveys in over 90 countries
- The project is implemented by ICF International.
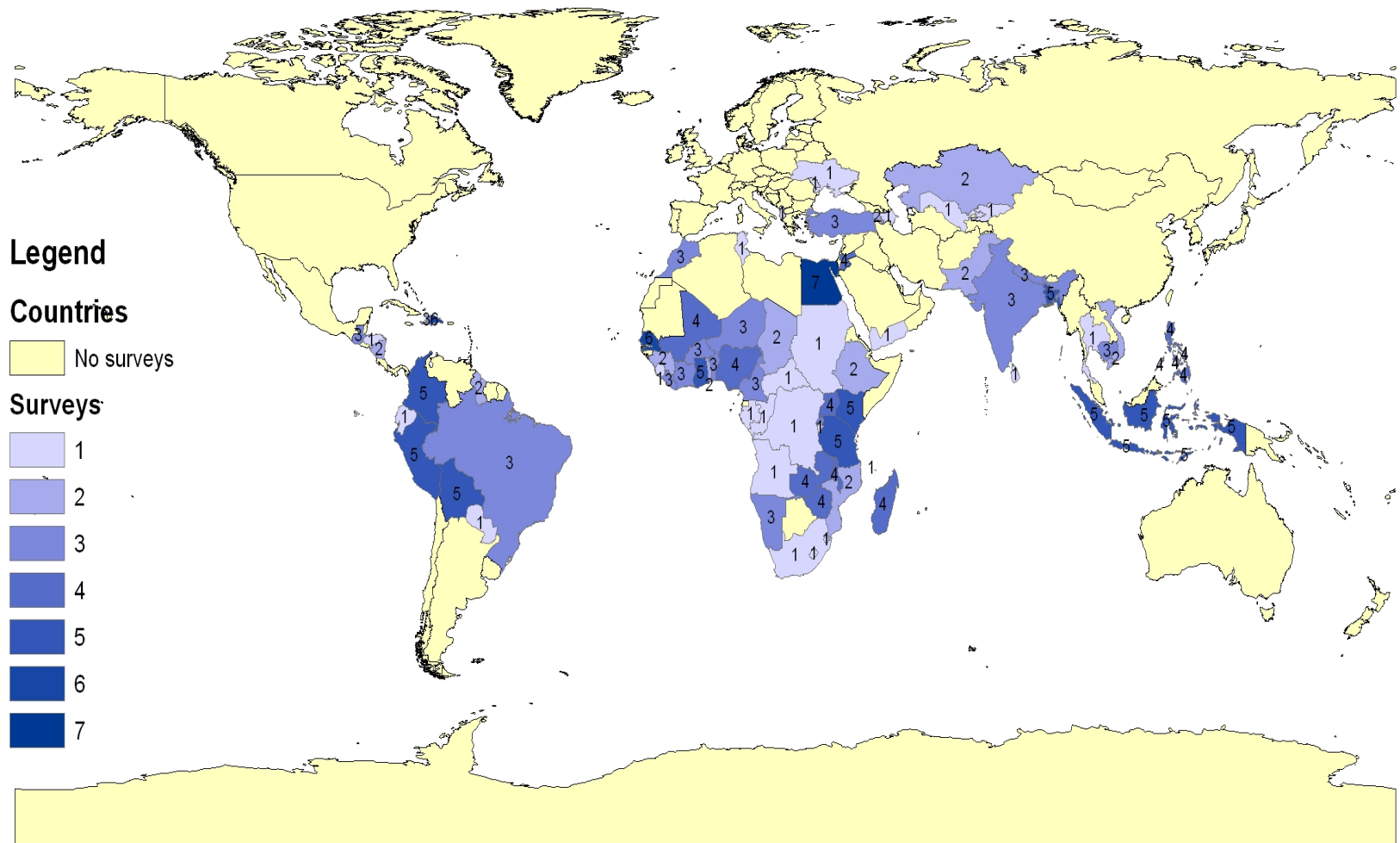- Primarily countries that receive USAID assistance

DHS Survey types

There are two main types of DHS Surveys:

- **Standard DHS Surveys:** large sample sizes (usually between 5,000 and 30,000 households) and typically conducted about every 5 years, to allow comparisons over time.

- **Interim DHS Surveys** focus on the collection of information on key performance monitoring indicators but may not include data for all impact evaluation measures (such as mortality rates)

- Other survey types:
    - AIDS Indicator Surveys (AIS)
    - Malaria Indicator Surveys (MIS)
    - Key Indicators Survey (KIS)
    - Service Provision Assessment (SPA) Surveys

DHS Survey topics

- Anemia
- Child Health
- Domestic Violence
- Education
- Environmental Health - water, sanitation, cooking fuel
- Family Planning, Fertility and Fertility Preferences
- Gender/Domestic Violence HIV/AIDS Knowledge, Attitudes, and Behavior
- HIV Prevalence
- Household and Respondent Characteristics Infant and Child Mortality
- Malaria
- Maternal Health and Maternal Mortality
- Nutrition Unmet Need
- Wealth
- Women's Empowerment
- Occupation of respondent and partner
- No information on income/expenditure

DHS regions

Legend

Countries

- No surveys

Surveys

- 1
- 2
- 3
- 4
- 5
- 6
- 7

Source: Günther and Harttgen 2013.

History of DHS and sample design

- **History of DHS**
  - DHS I: 1984–1989
  - DHS II: 1989–1993
  - DHS III: 1993–1997
  - MEASURE DHS+: 1997–2003
  - MEASURE DHS: 2003–2008
  - MEASURE DHS Phase III: 2008–2013

- **Sample Design**

  The sample is generally representative
  - At the national level
  - At the residence level (urban-rural)
  - At the regional level (departments, states)

  - In a majority of DHS surveys eligible individuals include women of reproductive age (15-49) and men age (15-59), or in some cases (15-54). In some countries only women are interviewed.

DHS Survey instruments

- **Questionnaires**: There are three core questionnaires in DHS surveys:
    1. A Household Questionnaire
    2. Women's Questionnaire
    3. Men's questionnaire
- B**iomarkers**
    - DHS surveys collect biomarker data relating to a wide range of conditions including infectious and sexually transmitted diseases, chronic illnesses such as diabetes and micronutrient deficiencies.
    - Most surveys now include testing for HIV infection in their survey design.

- **Geographic Information**
    - DHS surveys routinely collect geographic information in all surveyed countries.

- **Reports and Data**
    - Survey results for each country are published in a detailed final report
    - Data is made publicity available after the report has been published
    - Data can be downloaded from the DHS measure web page after registration

Example of Geographic Information Systems (GIS)

Defining urban slums

Slum Households in Addis Ababa, Ethiopia, 2005, UN-Habitat Definition

Source: Günther and Harttgen 2011.

DHS Data sets

- Data sets are available in several formats: SPSS, SAS, and STATA:
  - Individual Recode (IR)
  - Household Member Recode (PR)
  - Household recode (HR)
  - Children's Recode (KR)
  - Male Recode (MR)

UGPR60FL.MAP	Aug 28, 2012 6:42 PM	123 KB	Ausfü...–Datei
UGHR60FL.DTA	Aug 28, 2012 7:32 PM	36.5 MB	Stata Data File
UGIR60FL.DTA	Aug 28, 2012 7:32 PM	41.7 MB	Stata Data File
UGKR60FL.DTA	Aug 28, 2012 7:32 PM	8.8 MB	Stata Data File
UGMR60FL.DTA	Aug 28, 2012 7:32 PM	1.8 MB	Stata Data File

nadel-m02 ▸ Benutzer ▸ Kenneth ▸ Research ▸ Data ▸ DHS ▸ Uganda ▸ 2011

Survey Datasets: Household member recode and individual recode (Uganda 2011)

Household member recode

Data Editor (Browse) – UGPR60FL.DTA

hhid[1]   1 2

	hhid	hvidx	hv000	hv001	hv002	hv003	hv004	hv005	hv105
1	1 2	1	UG6	1	2	1	1	821594	25
2	1 2	2	UG6	1	2	1	1	821594	24
3	1 4	1	UG6	1	4	1	1	821594	41
4	1 4	2	UG6	1	4	1	1	821594	14
5	1 4	3	UG6	1	4	1	1	821594	13
6	1 9	1	UG6	1	9	1	1	821594	29
7	112	1	UG6	1	12	1	1	821594	20

Individual recode

Data Editor (Browse) – UGIR60FL.DTA

caseid[1]   1 4 1

	caseid	v000	v001	v002	v003	v004	v005	v012
1	1 4 1	UG6	1	4	1	1	885758	41
2	112 1	UG6	1	12	1	1	885758	20
3	121 2	UG6	1	21	2	1	885758	22

Survey data sets

- Matching data files

Unique Identifiers for Data Files

File	ID Variable	Cluster	HH Number	Line Number	Birth Order	Husband/ Wife
Household	HHID	HV001	HV002	HVIDX		
Women	CASEID	V001	V002	V003		V034
Men	MCASEID	MV001	MV002	MV003		MV034i
Children	CASEID	V001	V002	V003	BIDX	
Couples	CASEID	V001	V002	V003		

Source: DHS Guides to Statistics.

- Sampling Weights: Data must be weighted.
    gen weight=v005/1000

Web tools

- The MEASURE DHS website (www.measuredhs.com) allows free access to DHS data sets

- It also allows free access to survey publications, information on survey methodology and topics, survey status lists, and much more

- STATcompiler: online database tool that allows users to select from numerous countries and numerous population, health, and nutrition indicators to analyze trends across time and space. CAVEAT!! DATA ARE NOT CLEAN SO I WOULD NOT USE.

- Guide to DHS statistics:

Summary

- More than 200 surveys between 1984 and 2011
- Several rounds per country `(between 1 and 6)
- Pooled Data set: large sample size
- Rich information on household socioeconomic characteristics, household structure, family composition, geographic location
- Standardization of design (and coding)
- Allows analyzing changes in occupation
- No panel structure at individual level
- No information on income and/or expenditure
- No information on informal sector

What can DHS data tell about structural change?

- Occupational information from the Demographic and Health Surveys (DHS) may be used to document levels and changes in agriculture.

- Other occupational categories include: not working, professional/technical/managerial, clerical, sales, agricultural - self employed, agricultural – employee, household and domestic, services, skilled manual, unskilled manual and whether or not an individual is in school.

- Productivity cannot be measured.

- But it is possible to look at measures of wellbeing by occupation and changes in measures of wellbeing as occupational shifts occur.

- A similar exercise can be done

Distribution of occupations

	Total		Sub-Sahara Africa		Rest of the world	
	Respondent	Partner Mean	Respondent	Partner Mean	Respondent	Partner Mean
Occupation	Mean (%)	(%)	Mean (%)	(%)	Mean (%)	(%)
Not working	44.96	2.65	30.36	2.03	52.99	2.99
Professional/technical/managerial	4.9	11.01	2.56	8.07	6.19	12.62
Clerical	1.74	3.64	0.97	2.51	2.16	4.26
Sales	11.5	9.67	16.76	10.04	8.61	9.46
Agricultural - self employed	19.98	30.41	34.62	43.33	11.94	23.31
Agricultural - employee	5.15	8.09	3.87	8.09	5.85	8.09
Household and domestic	1.61	0.59	0.81	0.89	2.04	0.43
Services	3.08	6.77	2.86	5.59	3.21	7.42
Skilled manual	4.83	19.89	4.6	14.11	4.96	23.07
Unskilled manual	2.25	7.28	2.59	5.35	2.05	8.34
Total	100	100	100	100	100	100

Occupation groups by socioeconomic characteristics

	Respondent's occupation			
	Agriculture	Manual	Service	Not working
Total	22.90%	7.18%	22.88%	47.20%
By residence				
Rural	35.47%	6.17%	12.88%	45.48%
Urban	4.42%	8.34%	32.52%	54.72%
By household headship				
Female	20.60%	9.05%	27.82%	42.53%
Male	23.40%	6.55%	19.55%	50.50%
By education				
No education	35.32%	6.83%	12.75%	45.11%
Primary	30.11%	7.85%	16.20%	45.84%
Secondary or higher	7.08%	6.47%	31.59%	54.86%
By number of children				
0	15.70%	6.74%	19.42%	58.15%
3-Jan	21.79%	7.34%	23.51%	47.35%
4+	31.51%	6.75%	18.58%	43.16%
By Agegroup				
15-24	19.15%	6.04%	14.91%	59.90%
25-34	23.05%	7.65%	24.03%	45.27%
35-49	26.48%	7.49%	25.05%	40.97%

Summing Up Pros/Cons

- Pros
  - Wide coverage of very poor countries
  - Detailed measures of wellbeing
  - Gender, age, geographic location
- Cons
  - No measures of productivity
  - Occupation categories don't allow us to measure employment in manufacturing; best used to understand changes in agricultural employment shares.

Toward an Understanding of Economic Growth and Structural Change in Africa: Example from Tanzania

Key Messages

- Tanzania's recent growth is not well understood
- Significant positive structural changes (DHS, Census)
- But not "Asian" style (Ghana vs Vietnam)
- Tremendous growth in "In between" sector
- Need to understand "In between" sector's role
- Myriad of Industrial Policies targeted at this sector
- Little idea if any of these policies are working

Arthur Lewis: Part One

- 'An economy does not divide into a capitalist sector hiring workers for factories and other large units on the one hand, and a small farming sector on the other hand. <u>In between</u> are units of production of all sizes, and in particular a great number of one-to five-man undertakings in manufacturing, transport and a wide range of services – often nowadays called the informal sector. Some of this activity belongs in the modern sector as we have defined it; i.e., it will expand with economic development; the rest – e.g., some of the handicrafts and some of the services – belong to the traditional sector in that they will contract.

- Arthur Lewis, The Dual Economy Revisited, The Manchester School, V. 47, Issue 3, 1979 pp 211-299

Arthur Lewis: Part Two

- The expansion of small scale activity in the modern sector is an important part of the development process. This is not because it is a temporary resting-place for migrants from the countryside seeking jobs in large scale enterprise. In LDCs, no less than in MDCs (as we shall see in our next section) jobs in large scale urban enterprises are not normally awarded to people who have no connections. It is rather because this sector of the economy is useful in its own right, meeting genuine market needs, and providing a lot of employment in the process.'

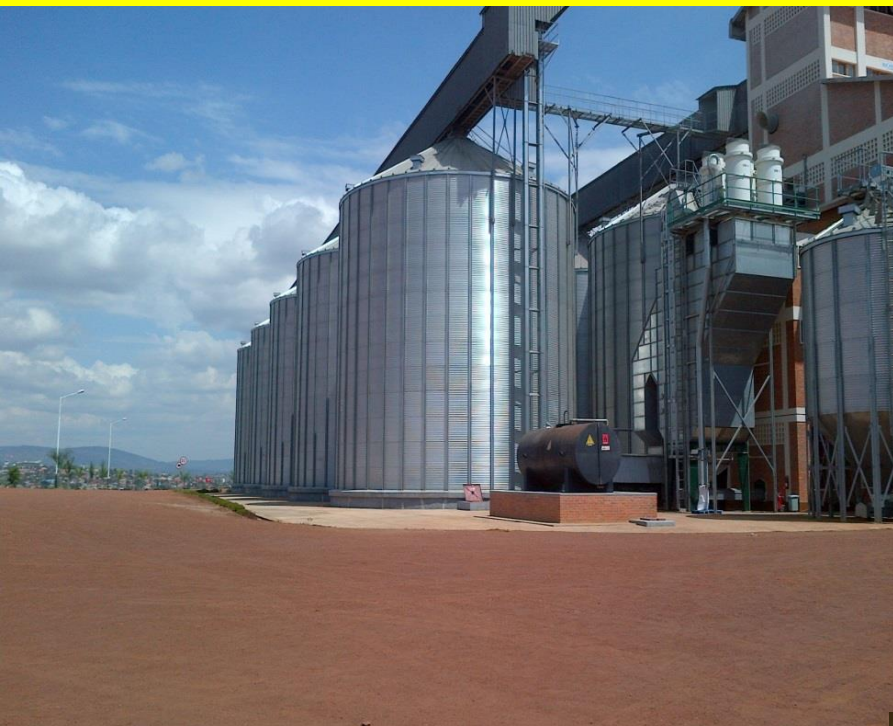- Arthur Lewis, The Dual Economy Revisited, The Manchester School, V. 47, Issue 3, 1979 pp 211-299

Why Don't We Understand the In Between Sector?

- Limited data – World Bank Enterprise Surveys
- Tendency to view ALL informal activities as backwards
- This is changing: Ethiopia, Nigeria, Tanzania, Rwanda
- First attempt to understand: we use Tanzania

Datasets Used

- Central Registry of Establishments (2005, 2010)

- Census (2002, 2012)

- National Panel Survey (2008, 2010, 2012)

- Formal Employment and Earnings Survey (2002-2013)

- MSME FIRST! Nationally Representative Survey (2010)

- Annual Survey of Industrial Production (Annual)

- World Bank Enterprise Survey (2006, 2013)

- Labor Force Surveys (2001, 2006)

Approach

1. Employment Trends by Sector, Firm Size and Formality Status

2. Productivity Trends by Sector, Firm Size and Formality Status

3. Drivers of the Growth in Small Firms: Public Services, Transportation Corridors

4. Potential for Small Firms to Contribute to Productivity Growth and Structural Change

Preliminary Results: Employment

Table 2: Formal Employment: age 15 and older, Mainland Tanzania 2002-2012

Formal Employment	2002	2012	Annual growth (%)	% of total growth	Notes
Total	14,934,140	16,976,933	1.3	100.0	Total Employment: Census data, >=15
Total (formal including micro firms)	1,007,074	1,780,143	5.9	100.0	Formal Employment: FEES data + CRE Guesstimate of micro firms in 2002 + 2012
Agriculture	101,670	98,001	-0.4	-0.6	
Mining and Quarrying	6,260	15,992	9.8	1.6	
Manufacturing	165,107	275,963	5.3	17.7	
Public Utilities	21,820	19,071	-1.3	-0.4	
Construction	46,203	46,237	0.0	0.0	
Commerce	139,591	350,947	9.7	33.7	
Transport and communication	46,344	67,496	3.8	3.4	
Finance, Community Services	187,510	188,778	0.1	0.2	

Source: Authors' calculations based on Formal Sector Employment & Earnings Reports (2007, 2011), National Bureau of Statistics, Tanzania, Census data (Minnesota Population Center (2014) and Census report (2014)), Central Register of Establishments (2005, 2010), National Bureau of Statistics, Tanzania.

Defining the In between Sector

Table 1: Definition of SMEs

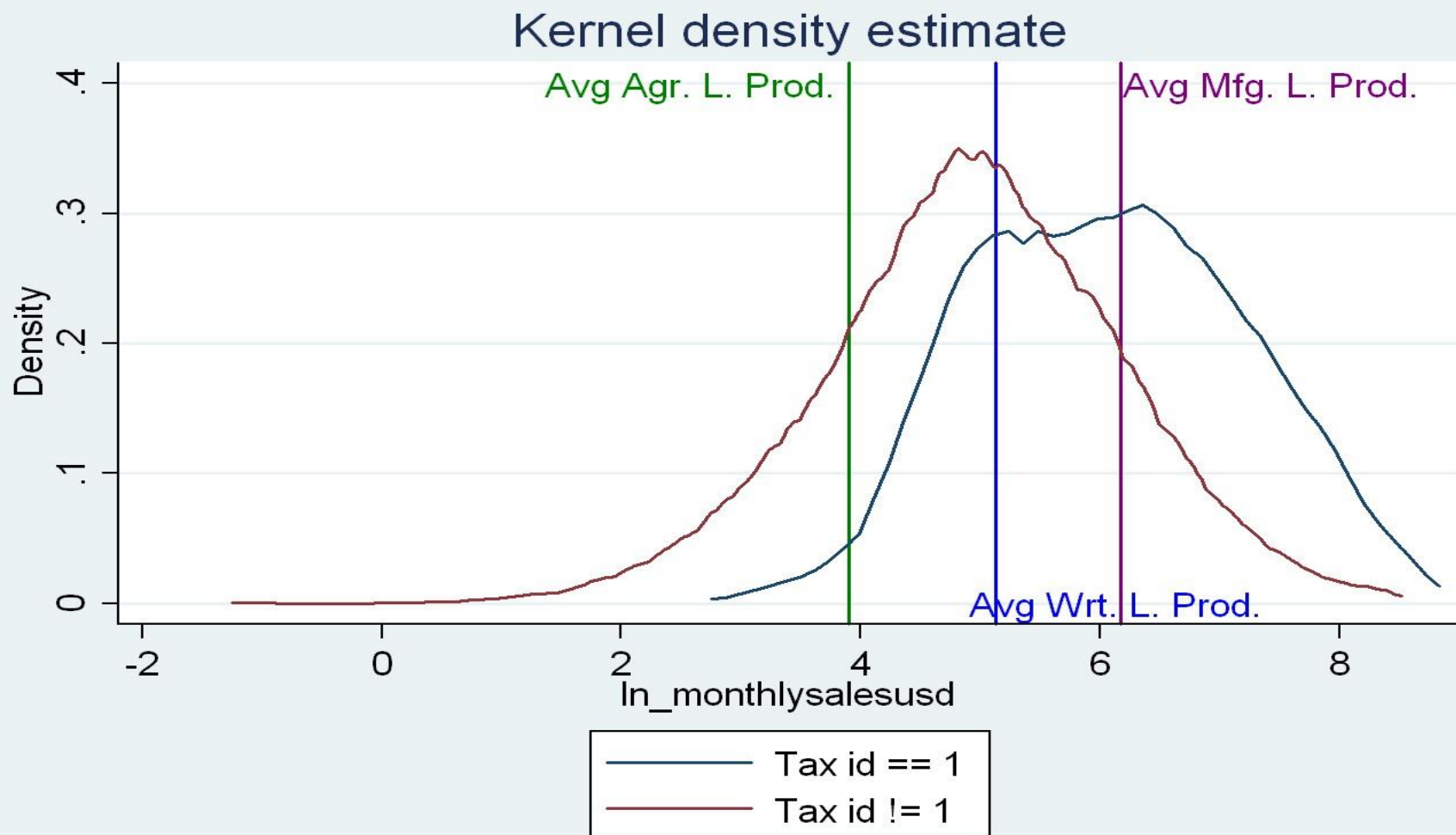| Category	Conceptual definition	Size definition (no. of employees)	Nature of Employment
Household Enterprises (HEs)	Own account Operators	1 or none	Informal, No paid employee (owner is same as operator)
Micro, Small and Medium Enterprises (MSMEs)	Micro Enterprises	1-4	Informal employees (without legal contracts)
	Small Enterprises	5 - 49	Mix of formal and informal Employees
	Medium Enterprises	50 - 99	Mostly formal employees
Corporate companies	Large Enterprises (LEs)	100 or more	Highly formal employees

Growth in In between Sector Example from Tanzania

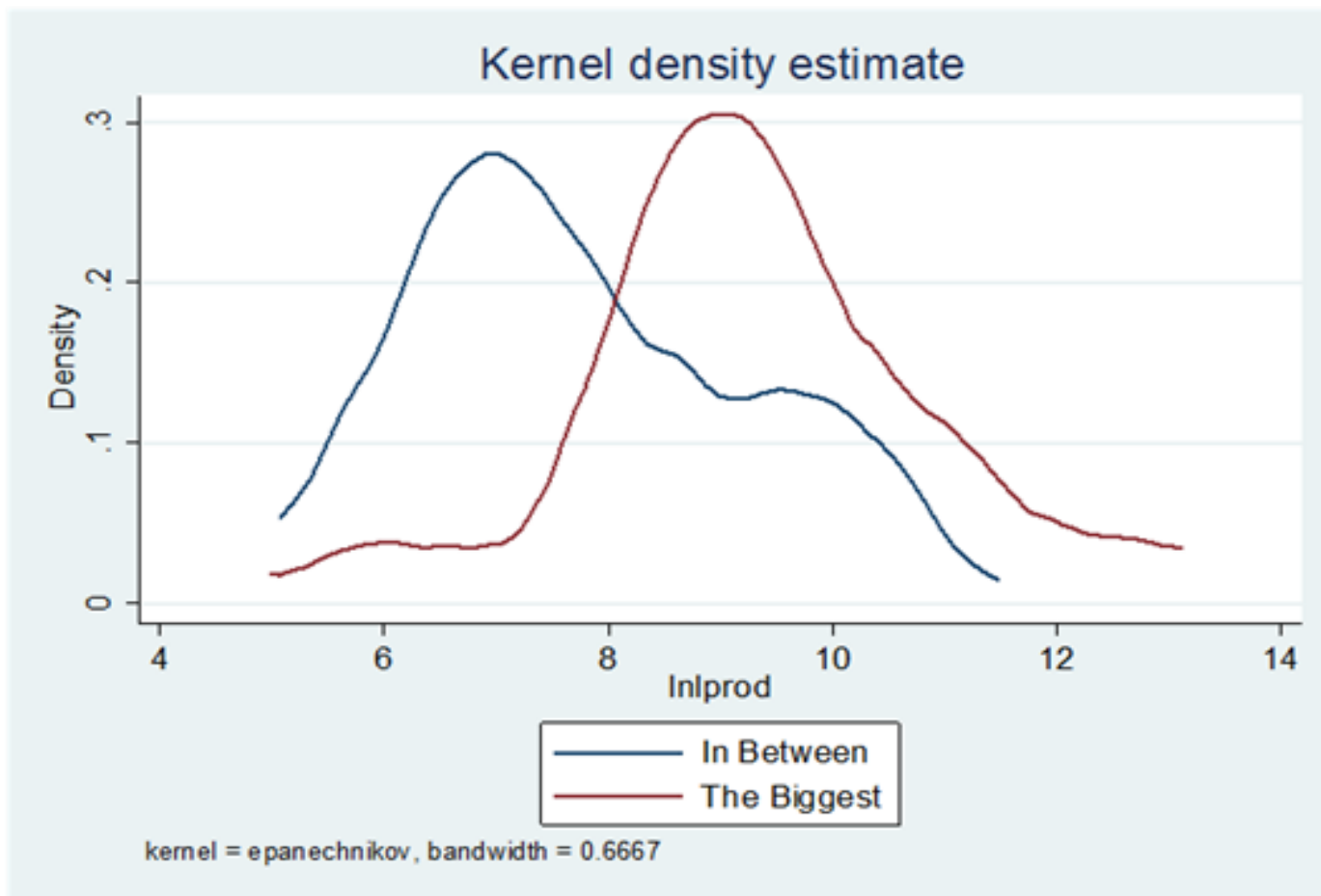**Table 3: Establishment distribution across employment categories in Tanzania**

Size Group	2005 Freq.	Percent	2009 Freq.	Percent	2010 Freq.	Percent
Micro [1-4]	19'282	47.13	28'395	34.64	31'589	37.01
Small [5-49]	18'518	45.26	51'161	62.41	51'451	60.28
Medium [50-99]	906	2.21	1'315	1.60	1'274	1.49
Large [100+]	974	2.38	1'111	1.36	1'042	1.22
n/a	1'231	3.01				
Total	40'911	100	81'982	100	85'356	100

Productive Heterogeneity: In Between Sector Tanzania MSME Survey 2012

kernel = epanechnikov, bandwidth = 0.3135

Productive Heterogeneity: In Between Sector Tanzania WBES Survey 2013

From the In Between Sector: Huajian China

FIRST BUSINESS START-UP, STARTING FROM NOTHING, HE FO

1984年10月在自己村庄创建青春鞋帽厂，发展拥有员工18人。
October 1984, Mr. Zhang Huarong set up Youth Shoes and Hat factory which had 18 workers in his village.

To the Modern Formal Sector: Huajian Ethiopia

Cautious Optimism: MSMEs Tanzania

- Tanzania's First Nationally Rep. SMSE Survey 2010
- More than 3 million such enterprises
- Employ around 5.2 million people
- 60 percent report businesses are growing;
- 54 percent of enterprise owners report that they would not give up their job for a full-time salaried position

- **Thus, while some of the enterprises belong in the traditional sector as conceptualized by Lewis (1954), many do not.**

Industrial Policy for Structural Change

- Facilitation of job creation in high productivity sectors
- Must include the in between sector
- Country strategies are extremely diverse
- 	- Ethiopia: Labor Intensive Manufacturing
- 	- Tanzania: Harnessing Natural Gas for Structural Change
- 	- Rwanda: Strategy Based on Knowledge Intensive Sectors
- But, 'In Between' sector important in <u>all</u> countries
- All countries have industrial policies vis a vis the ib sector
- So, not a question of to do or not to do
- A question of evaluation and targeting

Conclusion: Key Issues

- How do these sectors contribute to productivity growth? (Directly and Indirectly via Financing Public Investment)
- What kinds of policies are most effective for promoting job creation and productivity growth in the in between sector?
- Are some firms in the in between sector like Huajian eg do they have the potential to become large productive firms?