AMERICAN UNIVERSITY OF ARMENIA

CAPSTONE THESIS

Poverty Prediction with Alternative Data in the Context of Armenia

Author:

Aspram GRIGORYAN

Supervisor:

Dr. Hrant DAVTYAN

A thesis submitted in fulfillment of the requirements

for the degree of Bachelors

in the

BA in Business

College of Business and Economics

June 12, 2020

"*I thank everyone at AUA for my solid academic background, for the inspiration and the new perspective on life. I also thank my supervisor Dr Davtyan for his invaluable support in making this Capstone Thesis a reality and for conveying his passion towards the innovative alternative* "

Aspram

Abstract

Timely and accurate socioeconomic data are the cornerstone of effective policy formulation and evaluation afterwards. Over the years, the main source of such data in the developing world were household surveys and census which are subject to problems, including sampling and non-sampling errors, high cost and untimeliness of collection. As a solution, this paper avails an alternative source of data - satellite maps and imagery for poverty prediction in the context of Armenia. The findings prove, that a higher predictive power can be reached this way. Next, the employed model of Fixed effects of marz-level poverty prediction showcases that there is no significant marz-specific effect in poverty level, suggestive of no need of place-based policies in Armenia. The latter statement can be more accurately implied by future research that will involve a bigger scope of explanatory variables, possibly captured by alternative sources

Contents

List of Tables

Introduction

The elimination of worldwide poverty is the first of 17 UN Sustainable Development Goals for the year 2030. The World Social Summit identified poverty eradication as an ethical, social, political and economic imperative of mankind. To make progress to this end, governments worldwide are called on to address the root causes of it and to implement the necessary policies and actions to reduce its scale. To track the effectiveness of government policies and programs more frequent and reliable data on the distribution of poverty is needed.

The household surveys have been a major source of poverty pattern assessment worldwide. However, these data collection methods are error-prone, requiring considerable editing afterwards to clean the data. Also, these survey data are labour intensive and are typically costly to produce. Census data, on the other hand, is produced only with large intervals and is even more costly. As a solution a rising number of initiatives are focusing on harnessing the power of big data for socioeconomic analysis. In particular, satellite data have helped researchers to surpass many limitations of data-availability or data-collection. As the high-resolution spatial imagery becomes more and more available and represents a free alternative to traditional survey methods many researchers are underpinning their studies by availing this new source.

In this paper, I explore the possibility to overcome the shortcomings of standard sources in the context of Armenia through the combined power of satellite maps and imagery and Armenian national household survey data and statistical yearbooks. For the poverty assessment, the unit of analysis was chosen to be the marz (Yerevan included and referred to as a marz henceforth). The primary research question therefore, is “Does availing satellite imagery amount to a good proxy of marz level poverty indicator and a competitive alternative to traditional data sources in the context of Armenia". The approach combines remote sensing technology for data extraction and advanced panel data methods of Fixed and Random effects for analysis.

The findings show that the models employed are significant and the poverty levels can be predicted with 76% accuracy with the winner model. Based on the findings of the Fixed Effects model, I proceed to the discussion of "whether place-based policies should be implemented in addition to person-based policies for poverty eradication in Armenia". The disproportionalities of poverty dispersion and marz specific effects, if found, might indicate that place-based policies might be useful. If there is place and related contextual effect on marz poverty level then place-based economic policies, complementary to people-based policies,can influence the economic vitality of the marz.

Literature Review

Development Economics has long been concerned about what defines poverty and how it should be gauged. Today, as a result of ever-evolving research and studies there are multiple measurement methods including monetary and non monetary metrics, one-dimensional and multidimensional indices for poverty assessment. However, while numbers of high-resolution indicators of human welfare are routinely collected for populations in high-income countries, the geographical scattering of poverty in low- and middle-income countries is often uncertain. As a general rule, in the developing part of the world census data and household survey data are used to predict poverty levels across the population. This technique relies on the availability of data that is typically costly to produce and the collection process of which is error-prone. Recently, a number of studies have overcome this challenging procedure by availing alternative data sources.

The prior literature on this topic suggests that most of the existing poverty prediction alternative models are based on satellite night luminosity data. Night luminosity seems to be a good proxy of economic activity. For example, Mellander et al. show that luminosity level has a high correlation with the wage level in Sweden (R2=0.70) (2015). On the other hand, Henderson et al. show that lights data also allow for measurement of income growth in sub and supranational regions (2009). Noor et al. reveal that nightlight data correlate closely with asset-based indicators of wealth in thirty seven African countries. At intra-urban level, high clarity satellite imagery contribute with timely and accurate data about the space, opening up new avenues for exploring inequality dimension: the socio-spatial differentiation (Banzhaf et al., 2008). The later literature suggests that night-lights data is not sufficient for more in-depth analysis and higher accuracy. In particular, Jean et al. reveal that luminosity data alone has low performance in distinguishing poor and very poor rural areas in sub-Saharan Africa (2016). The shortcoming of luminosity maps, particularly in poor regions, has fueled papers that availed day-time high resolution satellite imagery (Jean et al, 2016) to measure poverty in developing countries. An interesting approach was effectively developed by Jean et al.(2016). They use a two step transfer machine learning framework .Their approach shows higher performance and an improved R2. Head et al successfully replicated the model used by Jean et al. Their research, however, reveals that satellite data is not effective at predicting a bigger range of human development indicators (2017). They also show that the performance of the algorithm is sensitive to the parameters used to train the convolutional neural network and that the same algorithm cannot be easily replicated to other geographical contexts or variables and it needs tuning. They imply that this approach of measuring poverty is not largely generalizable to other indicators and hardly implementable for “out-of-box” contexts.

In addition to overcoming the challenging process of poverty assessment through household surveys, the use of spatial imagery for poverty prediction implies that spatial and geographical factors of concentrated poverty will be counted for. Many studies on this topic suggest that areas of condensed poverty put extra burdens on poor households living in those areas, beyond what the family specific circumstances might imply. The research also indicates that areas of concentrated poverty can have wider effects on surrounding neighborhoods that are not classified as "high-poverty," thus limiting overall economic potential and social cohesion.

If there is concentrated poverty in a specific area of a country, then this might indicate the necessity to implement place-based policies. The economic phenomenon of labor mobility opposes place-based policies given that labor flows will arbitrage away spatial utility differences. (Ravallion and Wodon, 1999). This is the reason why economists argue against place policies including subsidies and tax cuts directed at poverty eradication in underprivileged communities. Therefore, they support person-based policies such as education programs, job counseling, and relocation assistance. In contrast to person-based policies there is a second thought which opposes limited labor mobility in low-income and middle-income countries or regions. It mainly argues that underprivileged households and workers with less human capital are not as geographically mobile (Ravallion and Wodon, 1999 2003).

Place-based policies derive their appeal from the notion that wide spatial variation in local attributes thwart "one-size-fits-all" policies. Place and related contextual effects influence economic vitality and shape the character of the people (Blank, 2004). In isolated inner cities and remote rural areas, many of the disadvantaged have less access to job training, counseling, healthcare, childcare, and transportation, suggesting that government-service delivery should reflect these spatial differences (Allard et al., 2003). Even in instances where person-based approaches may be appropriate, advocates of place-based policies argue they have an important complementary role (Blank, 2004). Place-based policy advocates also argue that economic development policies can effectively enhance local growth because of factors such as neighborhood effects, economic-role models, and knowledge spillovers. Finally, place-based policies have the simple advantage that governments may find it easier to target poor places than to identify households with the specific attributes that would merit targeting (Ravallion and Wodon, 1999).

Causal Relationship to be Studied

In the historical past, poverty was associated with a lack of shelter, food or water for an agent. As continued economic progress is at pace poverty has been losing its monotone meaning and quite often is considered to be a multidimensional figure. The Global Multidimensional Poverty Index (MPI) was developed in 2010 by the Oxford Poverty Human Development Initiative (OPHI) and the United Nations Development Programme and uses health, education and standard of living indicators to determine the degree of poverty experienced by the population.

Health	Child Mortality
	Nutrition
	Doctor visits per capita
Education	Years of schooling
	School attendance
Living Standards	Cooking fuel
	Sanitation
	Drinking Water
	Electricity
	Housing
	Assets

In this paper I view these dimensions of poverty as contributing factors for expenditure based poverty metric. I view poor health conditions and insufficient education as factors causing low levels of expenditure or as factors closely related to it. To capture these effects aggregated measures of household head education level and doctor visits per capita were used. In addition, nighttime luminosity based factor of standard of living was used. Luminosity is assumed to be correlated with asset based standard of living and naturally - the expenditure on electricity. Because asset-based measures of standard of living are regarded as a better proxy for the long-run household situation, luminosity data being correlated with it can be more representative of permanent income or long-run control of resources. This is important in the context of Armenia as a country heavily reliant on remittance flows from foreign countries. As opposed to consumption data that are highly noisy and capture short-run effects, asset-based metrics are better for long-term analysis. In addition luminosity data derive their appeal from the fact that they capture the whole population, all communities and households. This is in contrast with household survey asset-based standard of living metrics that are recorded only for a subset of the population. The sampling of the latter is troublesome and the sample might not perfectly reflect the population and produce misleading results.

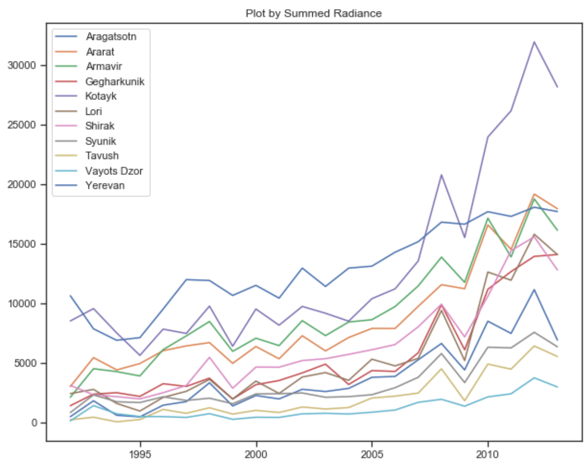
For the poverty level, marz-level average of weighted household expenditure was used. The already existing poverty categorization in the household surveys is based on the minimal consumption basket and is somewhat imperfect, because the minimal consumption basket is changing in the course of time and so does the threshold of poor, very poor and not poor. Therefore, we chose household expenditure - aggregated and averaged for each marz, as an indicator of poverty and standard of living.

Data Description

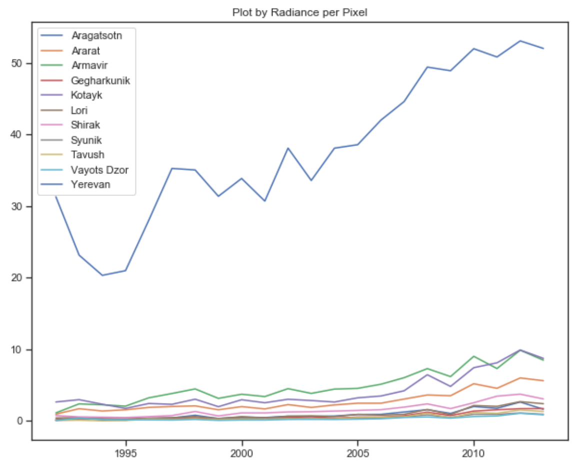
The constructs, respective metrics and the short description of sources and calculation of the variables included in the scope of this paper are represented in the table below.

Constructs	Metric Name	Source and Calculation
Poverty	average_expenditure	Weighted and averaged household expenditure for each marz and for a given year from hh survey data from households.
Health	hospital_visits_per_Capita	The hospital visits per capita for a given marz and for a given year, taken from the statistical yearbooks.
Education	headyears_educated	Numerified data on household head years spent in education from hh survey data from households.
Living Standards	summedradiance radianceoverpixels radpopscaled	Extracted from light-radiance maps, weighted over the pixels and population of a marz.
Other Variables	illegal_acts cargo_turnover passenger_turnover total_debt	Extracted from statistical yearbooks. Represent the crime rate, total credit and debit debt, and passenger and cargo turnover in a given marz for a given year.

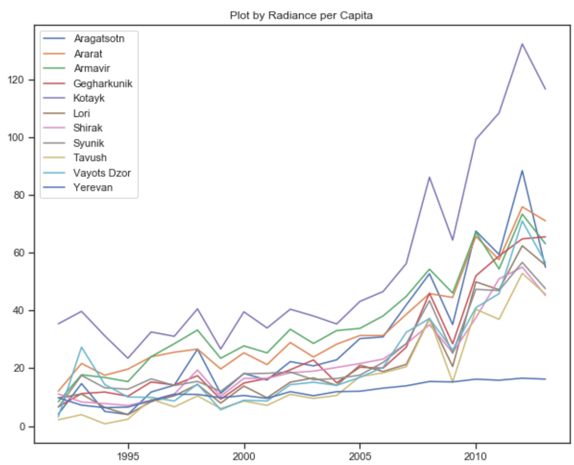
Nighttime luminosity data came from a light radiance emission map gathered by two satellite sensors. The data does not measure light-pollution or energy consumption directly, instead it is a radiance emission map that is assumed to have high correlation with the mentioned indicators. The data spans 1992-2013 and was captured by Defense Meteorological Satellite Program satellites. From 2012 to the present day the data comes from the Day/Night Band of the Visible Infrared Imaging Radiometer Suite instrument. Because of fundamental differences between the two, it is impossible to have a single dataset that will allow year-to-year comparisons from 1992-2020 using this particular source. This is the main reason why in the scope of this paper only a short time-span of 11 years was used (2004-2013). The other limitation that dictated the shortening of the time-span of the analysis was the household survey dataset and regional statistic publications by the RA statistical committee that were first carried out in 2004. Despite the splitting of the data for the model building, the whole radiance data was used for the descriptive plots below and showcases the dynamics from 1992 to 2013.



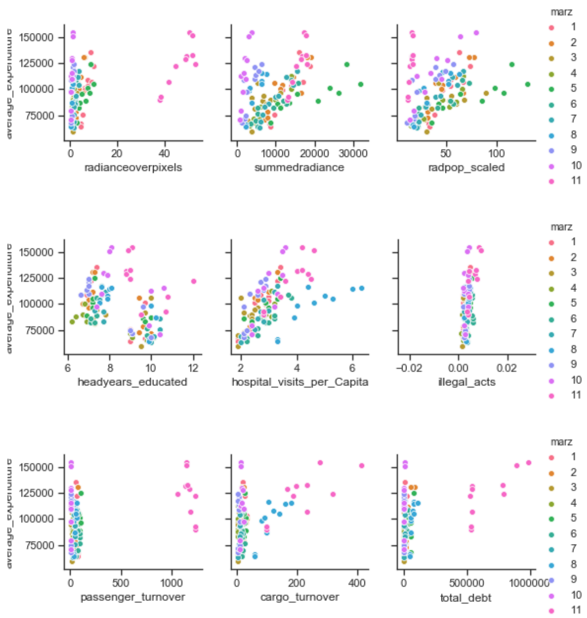
The plot by summed radiance across 10 marzes and the capital Yerevan captures the essential differences by marzes. Even though the dynamics are mostly coherent, there is a big gap between radiance emission levels across marzes. This is partially due to the bias of the indicator, as the summed radiance does not take account of the size of a marz. Radiance per pixel weights this indicator by the number of pixels that each marz pertains to in the map, however it only partially overcomes the problem, because a lot of uninhabited areas are taken into account. As the graph below indicates, this indicator creates a huge gap between the capital Yerevan and all other marzes.



The final transformation was calculating radiance per capita. As an individual level indicator it is assumed to be much more precise for poverty prediction. However, all the three indicators were included in the models as the other two are spatial factors that might show how isolated or condensed communities are in a certain marz, or in general, how large the marz is. This is in consistency with the poverty spatial concentration effect.



As for the household survey and the statistical yearbook publications, certain attributes were missing for the year 2004. These include hospital visits per capita, cargo and passenger turnover and total debt, that were filled by using backward filling, therefore the values of 2005 were used instead.



The scatter plots for the explanatory variables employed in the model and the dependent variable of average expenditure show that there is an identifiable pattern between summed radiance (summedradiance), per capita radiance(radpop_scaled), radiance over pixels measure (radianceoverpixels) and expenditure(average_expenditure) in marzes. The pattern is different for each marz as can be seen by different colors that correspond to different marzes.

Similarly, hospital visits per capita (hospital_visits_per_Capita) seem to have an increasing trend over time. Some other indicators that were extracted from statistical yearbooks, such as the total cargo turnover (cargo_turnover), passenger turnover (passenger_turnover), total debit and credit debts (total_debt) and crime level (illegal_acts) do not show much variance over time but the grouping by marzes is still evident.

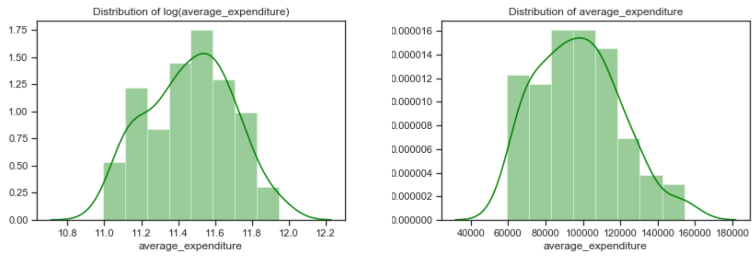
Regarding the averaged years educated, the data has been digitized. In the household surveys the household head education level was presented by categories. The following values were given to each category - 1 point corresponding to 1 year of education.

N	Value Labels	Numerical equivalents
1	illiterate	0
2	uncompleted primary	2
3	primary	3
4	uncompleted low secondary	5
5	low secondary	6
6	uncompleted upper secondary	8
7	upper secondary	10
8	secondary specialized	11
9	uncompleted high	12
10	high education	14
11	post-graduate	16

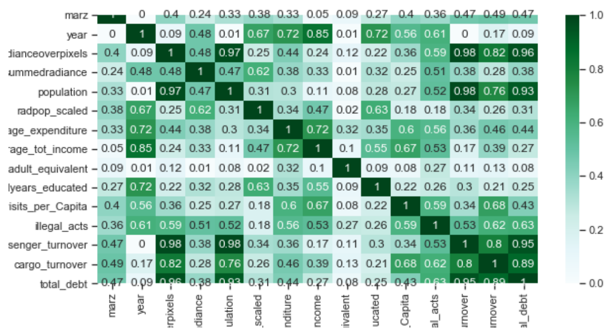
The indicator itself, however, is controversial because while marz differences are expected, no drastic changes are expected in average years of educated over the timespan of ten years. Also, the data processing showed that there were some unexpected ups and downs in average years educated in certain marzes for certain years. The scatter plot shows the unexpected jump of education level over time. Because different households are surveyed each year, even though the household types are weighted, the process is very error prone and it generates figures that are quite imperfect for time series analysis. The shortage of reliable marz level data once again accentuates the importance of using alternative sources for assessing the poverty levels.

Estimation and Hypothesis Testing

For model estimation linear panel models were used. Both dependent and independent variables have been log-transformed to reduce the skewness.



The heatmap below shows the correlation between all the variables gathered. Some of the variables exhibit high correlation. The highly correlated variables and also variables showing little to no time variance were excluded and the models were regressed over a few remaining variables.



Given that the models are multilevel and individual specific effects are assumed to exist, the base model of Pooled OLS cannot solve the problem. Instead, Random Effects model was employed. This is in consistency with the fact that the data comprises the same entities (marzes) observed each year. Therefore, even though the Pooled OLS model shows a high adjusted R^2 of 73% (Table 9.1, Appendix), it should not be employed.

The output table of the Random Effects model(Table 9.2, Appendix) shows that average expenditure was regressed over the three radiance indicators, education level indicator (headyears_educated), health indicator (hospital_visits_per_Capita) and crime rate indicator (illegal_acts). An adjusted R^2 of 74,6% was reached with this model. As an extension of the OLS model, Random Effects model includes the constant in the vector of regressors, and the error is being composed by both an unobserved effect (time invariant), and an observed error (time variant). The underlying random effects framework assumption here is that the unobserved effect is uncorrelated with the explanatory variables.

When this assumption is not satisfied, Fixed Effects model should be used instead. The FE model allows for an arbitrary dependence between omitted effect and the observed explanatory effects - leading to individual fixed effect.

The Fixed Effects regression model can be represented by the following equation

$$Y_{it} = \beta_1 X_{1,it} + \cdots + \beta_k X_{k,it} + marz_i + u_{i,t}$$

where *i= 1,...,11 and t=1,..., 10* and *i* are the entity specific intercepts that capture heterogeneities across marzes.

To understand whether random effects or fixed effects assumption holds, Hausman test was carried.

Hausman Test

data: log(average_expenditure)   log(radianceoverpixels) + log(radpop_scaled) + ...
chisq = 75.295, df = 6, p-value = 3.337e-14 alternative hypothesis: *one model is inconsistent*

The p-value of 3.337e-14 indicates of the fact that Fixed Effects model should be used.

Four different Fixed Effect models have been used, corresponding to different experiments with the variables. The first model contains only variables from hh survey data. The second model includes all the variables, the third includes only satellite data and variables that were not correlated with those. Finally the fourth model includes only variables from alternative data sources.

TABLE 5.1: Linear Panel Fixed Effects Models of Average Expenditure

	Dependent variable:			
	log(average_expenditure)			
	(1)	(2)	(3)	(4)
log(radianceoverpixels)		0.252*	0.234**	0.056
		(0.147)	(0.118)	(0.350)
log(radpop_scaled)		0.192	0.209	1.555**
		(0.666)	(0.611)	(0.600)
log(summedradiance)		−0.283	−0.284	−1.287*
		(0.620)	(0.594)	(0.735)
log(headyears_educated)	−0.040	0.100	0.105	
	(0.139)	(0.143)	(0.139)	
log(construction_mln_amd)		0.007		
		(0.014)		
log(hospital_visits_per_Capita)	1.126***	0.806***	0.836***	
	(0.132)	(0.202)	(0.159)	
log(illegal_acts)	−0.105*	−0.082	−0.081*	
	(0.057)	(0.059)	(0.047)	
log(total_credit_debt + total_debitory_debt)		0.004		
		(0.032)		
log(passenger_turnover)		−0.020		
		(0.086)		
log(cargo_turnover)		0.010		
		(0.042)		
Observations	110	110	110	110
R^2	0.760	0.799	0.798	0.698
Adjusted R^2	0.728	0.754	0.764	0.657
F Statistic	101.360*** (df = 3; 96)	35.385*** (df = 10; 89)	61.348*** (df = 6; 93)	74.080*** (df = 3; 96)
Note:				*p<0.1; **p<0.05; ***p<0.01

The table showcases the four models. All the four models are significant. The highest adjusted R^2 is achieved when the variables corresponding to passenger and cargo turnover and total debt that have priorly found to be correlated with radianceoverpixels, are eliminated. This model has an adjusted R^2 of 76.4%. The 1st and 2nd models which were created to compare the information added by the multidimensional poverty index components (1st) and possible explanatory variables from survey data perform less better than model 3. In addition, the model that consists of only satellite data achieves an adjusted R^2 of 65.7%.

All the four models in the table eliminate the omitted bias caused by excluding unobserved variables that differ across entities but are constant over time. The table below combines three models on the same variables but with different fixed effects. The first represents a model with time effects, it eliminates the bias from unobservables that change over time but are constant across entities. The second corresponds to the initial model with entity fixed effects and the last is the combined, two-ways model. The combined model allows to eliminate bias from unobservables that change over time but are constant over entities as well as it controls for factors that differ across entities but are constant over time.

TABLE 5.2: Linear Panel Fixed Effects Models Including Entity, Time and Twoways Effects

	Dependent variable:		
	log(average_expenditure)		
	(1)	(2)	(3)
log(radianceoverpixels)	0.234**	0.164***	0.185
	(0.118)	(0.017)	(0.116)
log(radpop_scaled)	0.209	0.078	0.296
	(0.611)	(0.051)	(0.673)
log(summedradiance)	−0.284	−0.249***	−0.391
	(0.594)	(0.058)	(0.719)
log(headyears_educated)	0.105	−0.016	0.501**
	(0.139)	(0.241)	(0.231)
log(hospital_visits_per_Capita)	0.836***	0.210***	0.088
	(0.159)	(0.070)	(0.245)
log(illegal_acts)	−0.081*	−0.021	−0.032
	(0.047)	(0.055)	(0.068)
Observations	110	110	110
R^2	0.798	0.629	0.083
Adjusted R^2	0.764	0.570	−0.190
F Statistic	61.348*** (df = 6; 93)	26.563*** (df = 6; 94)	1.265 (df = 6; 84)
Note:			*p<0.1; **p<0.05; ***p<0.01

The F Statistic of overall significance shows that the combined, two-ways model does not provide a better fit to the data than a model that contains no independent variables. Because the model that employs time-effects only and the combined model estimate the relationship quite differently and the coefficients change significantly, this means that the estimated relationship is affected by omitted variable bias due to factors that are constant over time but different across entities. The table below represents the entity fixed effects in deviation from the overall intercept, their standard errors and the test of equality to the overall intercept.

TABLE 5.3: Entity Fixed Effects For Marzes

	Estimate	Std. Error	t-value	Pr(>|t|)
1	-0.071	3.890	-0.018	0.985
2	0.075	3.970	0.019	0.985
3	0.049	3.734	0.013	0.990
4	0.112	4.084	0.027	0.978
5	-0.057	3.971	-0.014	0.989
6	0.038	4.045	0.009	0.992
7	0.069	4.015	0.017	0.986
8	-0.248	3.845	-0.064	0.949
9	0.209	3.696	0.056	0.955
10	-0.049	3.330	-0.015	0.988
11	-0.127	4.296	-0.030	0.976

The test of equality to the overall intercept shows that the entity effects do not deviate significantly from the overall intercept.

Discussion of the Results

The findings show that the usage of satellite data adds information and improves the model predictive power when combined with traditional source data (improving the adjusted R^2 from 72,8% to 76,4%) and when applied separately, represents a competitive alternative (Adj. $R^2 = 65{,}7\%$). This is in coherence with the first research question and research expectations of getting an improved poverty assessment model through the use of alternative data. Because only a single satellite metric of radiance was used with three different transformations to capture the population and the area of the marz, this means that the future studies might reach even higher predictive power by capturing complementary correlates of poverty levels and behaviour. For example data of physical properties, such as rainfall, temperature and vegetation capture information related to agricultural productivity, while distance to roads and cities, the dynamics of infrastructure reflect access to markets and information. In addition, timely credit consumption data on mobile phones and the number of agents in a marz using mobile phones show the inclusion level of households to use financial resources, while the spatial reach of the calling networks of household members may be indicative of remittance flows and economic opportunities. Integration of all possible alternative data has a potential of providing real-time, more accurate and less costly prediction. Moreover, both Random effects and Fixed effects models were used for poverty prediction, but the Hausman test showed that Fixed Effects should be used. Therefore, fixed effects were computed by showing no significant deviation from the overall intercept. This suggests having no significant marz-specific effects and might be talking about no necessity of implementing place-based policies of poverty eradication. The collection of additional data and the inclusion of some other important explanatory variables might not only increase the accuracy of the models and the predictive power but also be important in finding out whether there is a need of place-based policies complementary to people-based policies.

Conclusion

The elimination of poverty is an imperative for Armenia. Despite satisfactory economic growth indicators, poverty levels remain high amounting to 20% of the population. Accurate poverty tracking and assessment is an important factor in developing effective policies for poverty eradication, however, the traditional data sources of household survey data and statistical yearbooks are a costly, time consuming and error-prone tool for socioeconomic analysis. The use of alternative data has shown significant improvements in many poverty studies and this paper scales that method for Armenia, significantly improving the predictive power of the models employed. The inclusion of other alternative data in future analysis have the potential to integrate a wide range of important factors such as the dependency of communities from weather conditions, vegetation level or rainfall, the infrastructure dynamics, financial inclusion level and others that can accurately be extracted from non-traditional sources. A more inclusive model can also help to more certainly understand whether place-based policies are appropriate in the context of Armenia.

References:

Banzhaf, E., Hofer, R. (2008). Monitoring Urban Structure Types as Spatial Indicators With CIR Aerial Photographs for a More Effective Urban Environmental Management. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 1(2), 129–138. doi: 10.1109/jstars.2008.2003310

Blank, R. "Poverty, Policy and Place: How Poverty and Policies to Alleviate Poverty Are Shaped by Local Characteristics." Rural Poverty Re- search Center Working Paper 04-02, Corvallis, OR, 2004.

Chambers R. 2006What is poverty? Who asks? Who answers? Poverty in Focus December 2006, UNDP International Poverty Centre.
See http://opendocs.ids.ac.uk/opendocs/handle/123456789/120. Google Scholar

Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A. (2014). Return of the Devil in the Details: Delving Deep into Convolutional Nets. Proceedings of the British Machine Vision Conference 2014. doi: 10.5244/c.28.6

Deaton, A. (2019). Front Matter. The Analysis of Household Surveys (Reissue Edition with a New Preface): A Microeconometric Approach to Development Policy, i-xiv. doi: 10.1596/978-1-4648-1331-3_fm

Donaldson, D., Storeygard, A. (2016). The View from Above: Applications of Satellite Data in Economics. Journal of Economic Perspectives, 30(4), 171–198. doi: 10.1257/jep.30.4.171

Jargowsky, P.A. "Stunning Progress, Hidden Prob- lems: The Dramatic Decline of Concentrated Poverty in the 1990s." Living Cities Census Se- ries. Washington DC: The Brookings Institu- tion, 2003

Head, A. E., Manguin, M. E., Tran, N. E., Blumenstock, J. E. (2017). Can Human Development be Measured with Satellite Imagery? Proceedings of the Ninth International Conference on Information and Communication Technologies and Development - ICTD 17. doi: 10.1145/3136560.3136576

Henderson, J. V., Storeygard, A., Weil, D. (2009). Measuring Economic Growth from Outer Space. doi: 10.3386/w15199

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. Science, 353(6301), 790–794. doi: 10.1126/science.aaf7894

Mellander, C., Lobo, J., Stolarick, K., Matheson, Z. (2015). Night-Time Light Data: A Good Proxy Measure for Economic Activity? Plos One, 10(10). doi: 10.1371/journal.pone.0139779

Michael Xie, Neal Jean, Marshall Burke, David Lobell, and Stefano Ermon. (2016). Transfer learning from deep features for remote sensing and poverty mapping. AAAI Conference on Artificial Intelligence, arXiv:1510.00098v2 (Feb. 2016). https: //arxiv.org/abs/1510.00098v2

Partridge, Mark Rickman, Dan. (2005). Persistent pockets of extreme american poverty: People or place based?.

Poverty eradication | UN. (n.d.). Retrieved June 12, 2020, from https://www.un.org/development/desa/so
eradication.html

Poverty reduction | UNDP in Armenia. (n.d.). Retrieved June 12, 2020, from https://www.am.undp.org/content/armenia/en/home/sustainable-development.html

Ravallion, M. and Wodon, Q. (1999), Poor Areas, or Only Poor People?. Journal of Regional Science, 39: 689-711. doi:10.1111/0022-4146.00156

Steele JE et al. 2017. Mapping poverty using mobile phone and satellite data. J. R. Soc. Interface 14: 20160690. http://dx.doi.org/10.1098/rsif.2016.0690

Appendix

TABLE 9.1: Pooled OLS model

	Dependent variable:
	log(average_expenditure)
log(radianceoverpixels)	0.197***
	(0.020)
log(radpop_scaled)	0.210***
	(0.057)
log(summedradiance)	−0.300***
	(0.056)
log(headyears_educated)	−0.292***
	(0.105)
log(hospital_visits_per_Capita)	0.451***
	(0.058)
log(illegal_acts)	−0.014
	(0.043)
Constant	13.451***
	(0.415)
Observations	110
R^2	0.747
Adjusted R^2	0.732
F Statistic	50.688*** (df = 6; 103)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE 9.2: Random Effects Model

	Dependent variable:
	log(average_expenditure)
log(radianceoverpixels)	0.197***
	(0.031)
log(summedradiance)	−0.265***
	(0.073)
log(radpop_scaled)	0.229***
	(0.065)
log(headyears_educated)	−0.100
	(0.106)
log(hospital_visits_per_Capita)	0.567***
	(0.075)
log(illegal_acts)	−0.046
	(0.044)
Constant	12.754***
	(0.522)
Observations	110
R^2	0.760
Adjusted R^2	0.746
F Statistic	326.836***
Note:	*p<0.1; **p<0.05; ***p<0.01