Measuring sustainable development progress in Peru using a multivariate latent Markov model.

F. Roscioli[1] D. Malerba[2] A. Farcomeni[1]

[1] *University of Rome Tor Vergata*
[2] *German Institute of Development and Sustainability (IDOS)*

6th September 2024

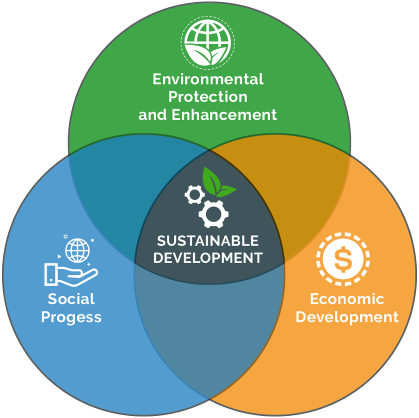
- **Development** is a complex and multifaceted phenomenon that encompasses profound transformations in social structures, attitudes, and institutional frameworks, in addition to economic expansion, mitigation of inequality, and elimination of absolute poverty (Todaro & Smith, 2009).

- **Development** is a complex and multifaceted phenomenon that encompasses profound transformations in social structures, attitudes, and institutional frameworks, in addition to economic expansion, mitigation of inequality, and elimination of absolute poverty (Todaro & Smith, 2009).
- **Sustainable development** (SD) means **development that can be continued** either indefinitely or for a given time horizon (Lélé, 1991).

- **Development** is a complex and multifaceted phenomenon that encompasses profound transformations in social structures, attitudes, and institutional frameworks, in addition to economic expansion, mitigation of inequality, and elimination of absolute poverty (Todaro & Smith, 2009).
- **Sustainable development** (SD) means **development that can be continued** either indefinitely or for a given time horizon (Lélé, 1991).
- This implies that the best choices are likely to remain those that meet the needs of society and are **environmentally and economically viable**, **economically and socially equitable** as well as **socially and environmentally bearable** (Porter & Linde, 1995).

- **Development** is a complex and multifaceted phenomenon that encompasses profound transformations in social structures, attitudes, and institutional frameworks, in addition to economic expansion, mitigation of inequality, and elimination of absolute poverty (Todaro & Smith, 2009).
- **Sustainable development** (SD) means **development that can be continued** either indefinitely or for a given time horizon (Lélé, 1991).
- This implies that the best choices are likely to remain those that meet the needs of society and are **environmentally and economically viable**, **economically and socially equitable** as well as **socially and environmentally bearable** (Porter & Linde, 1995).
- The highlighted complexity clearly shows how **measuring SD remains a challenging task**.

Sustainable Development



Measurement

"We can't manage what we cannot measure."
Peter Drucker

Measurement

"We can't manage what we cannot measure."
Peter Drucker

- Gross Domestic Product

Measurement

"We can't manage what we cannot measure."
Peter Drucker

- Gross Domestic Product
- Multidimensional Poverty Index (UNDP-OPHI)

Measurement

"We can't manage what we cannot measure."
Peter Drucker

- Gross Domestic Product
- Multidimensional Poverty Index (UNDP-OPHI)
- Human Development Index (UNDP)

Measurement

"We can't manage what we cannot measure."
Peter Drucker

- Gross Domestic Product
- Multidimensional Poverty Index (UNDP-OPHI)
- Human Development Index (UNDP)
- Sustainable Development Goals (UN)

Novelty

- Using a **multivariate latent Markov model** (Bartolucci et al., 2013), we aim at building up an alternative measure of sustainabile development.
- We apply our model to microdata from **Peru 2004-2017**: gdp increased on average by 5% annually (World Bank, 2024), the poverty headcount was reduced from 16.3% to 4.5% and inequality measured by Gini index decreased by 6.6 percentage points (World Bank, 2024) (INEI, 2018).
- The novelty of our approach allows us to address several critics in the literature on composite indicators for the measurement of development by suggesting a different methodological approach with a similar theoretical framework.

Novelty

- Using a **multivariate latent Markov model** (Bartolucci et al., 2013), we aim at building up an alternative measure of sustainabile development.
- We apply our model to microdata from **Peru 2004-2017**: gdp increased on average by 5% annually (World Bank, 2024), the poverty headcount was reduced from 16.3% to 4.5% and inequality measured by Gini index decreased by 6.6 percentage points (World Bank, 2024) (INEI, 2018).
- The novelty of our approach allows us to address several critics in the literature on composite indicators for the measurement of development by suggesting a different methodological approach with a similar theoretical framework.

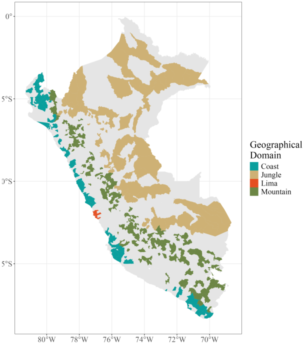
Research Questions

- **Can we measure sustainable development during the considered period using our model?**
- **If so, is our measure correlated with economic development?**

Data Sources

Balanced panel of 528 districts.

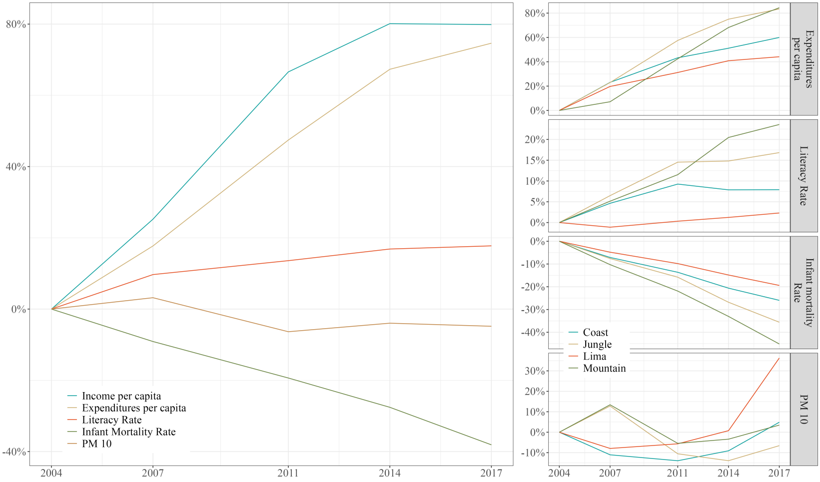
- **Consumption per capita** and **literacy rate**: *Encuesta de Hogares*, cross-sectional data from waves 2004, 2007, 2011, 2014 and 2017.
- **Infant mortality rate at birth** (INEI) by district (2007, 2017).
- **PM_{10} concentration** from the ECMWF Atmospheric Composition Reanalysis 4 (Inness et al., 2019) expressed in $\mu g/m^3$, and normalized by the annual reference average concentration of 15 $\mu g/m^3$ set by the World Health Organization (WHO)(Organization et al., 2021). Resolution of 0.75° x 0.75° (about 27 square kilometers). Air Pollution



Measurement comparison

Name	Measure	Dimension	Weighting	Method	Unit of measure
GDP	GDP	Economic	-	Single indicator	USD
HDI	GNI	Economic	1/3	Composite indicator	Score
	Education Index	Social	1/3		
	Life Expectancy	Social	1/3		
Full model	Income	Economic	Data Driven	LMM	Latent Classes
	Literacy Rate	Social			
	Infant Mortality Rate	Social			
	PM_{10}	Environmental			

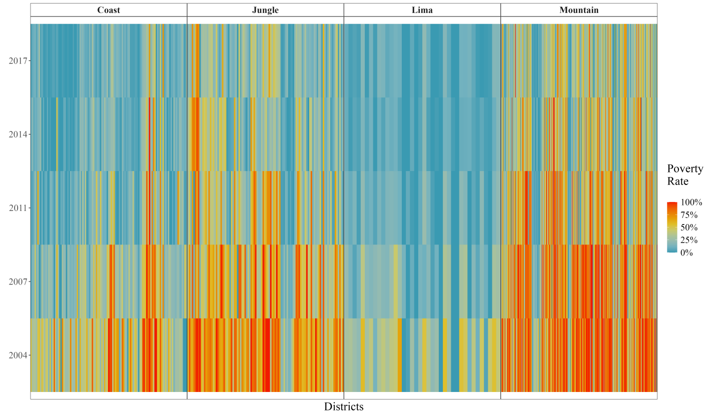
Descriptives statistics



Percentage change in the annual median value of the selected variables with respect to 2004. Left panel at national level; Right panel by geographical domain.

Variability

Income Poverty



Share of household income poor by district and year in Peru. The shade is the poverty rate in each district as defined by INEI. On the horizontal axes, there are districts clustered by geographical domain; on the vertical axes, the years considered. A change in color from red to cyan in the vertical direction for a given district implies reduction of the poverty rate in that district.

The model

$$E[Y_{itj}|U_{it} = c, X_{it}] = \alpha_{cj} + \beta_j X_{it}$$

- We assume that the results are independent of each other and of the past conditionally on a discrete latent variable U_{it} that can take values from 1 to $\boldsymbol{k}$.
- We also have a column vector of covariates X_{it} for each district on each occasion of time (time dummies).

The model

$$E[Y_{itj}|U_{it} = c, X_{it}] = \alpha_{cj} + \beta_j X_{it}$$

- We assume that the results are independent of each other and of the past conditionally on a discrete latent variable U_{it} that can take values from 1 to $\boldsymbol{k}$.
- We also have a column vector of covariates X_{it} for each district on each occasion of time (time dummies).
- Let Y_{itj} denote the j-th Gaussian response, for household $i = 1, \ldots, n$ at time $t = 1, \ldots, T_i$.
- Y_{it1} to denote expenditures per capita (logs).
- Y_{it2} for literacy rate (logs of the odds ratio).
- Y_{it3} for infant mortality rate (logs of the odds ratio).
- Y_{it4} for PM_{10} (logs).

HDI model

Class	Expenditures Per Capita	Literacy Rate	Infant Mortality Rate
1	23.3***	0.56***	0.34***
2	**68.49*** **	**2.21*** **	**0.16*** **
3	47.47***	1.29***	0.29***
4	**26.8*** **	**0.71*** **	**0.56*** **

Table: Intercept parameters in the HDI model. Expenditures per capita are expressed in *Peruvian Soles*, literacy rate and infant mortality rate can be instead interpreted as relative risks. In red we hilighted what we consider the "worse" class, and in cyan, the "best" class. ., * , ** , *** Significance at the 10, 5, 1and 0,1% levels, respectively.

HDI model

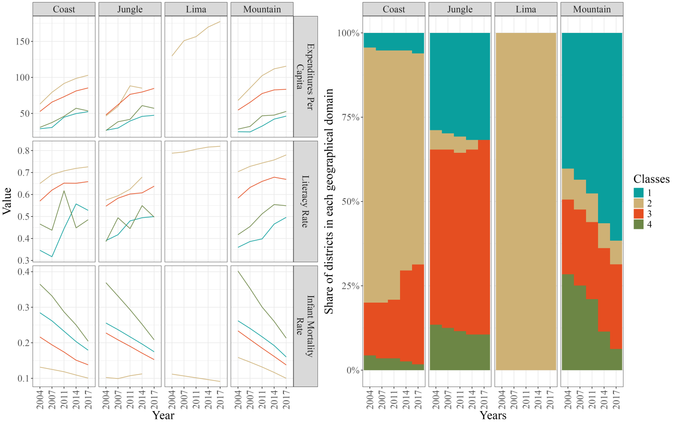


Figure: Left panel: average value of the considered variables by HDI model class, year and geographical domain. Right panel: the share of districts present in each HDI model class by geographical domain.

Full model

Class	Expenditures Per Capita	Literacy Rate	Infant Mortality Rate	PM_{10}
1	59.51***	1.93***	0.2***	0.68***
2	27.01***	0.76***	0.44***	0.56***
3	**68.17*****	**2.08*****	**0.19*****	**1.5*****
4	**28.3*****	**0.67*****	**0.35*****	**0.98.**

Table: Intercept parameters in the full model. To each class corresponds a value in each response variable. Income is expressed in *Peruvian Soles*; literacy rate and infant mortality in odds ratio; PM_{10} is normalized by the reference value of $15\mu g/m^3$. In red we highlighted what we consider the "worse" class, and in cyan, the "best" class. ., * , ** , *** Significance at the 10, 5, 1and 0,1% levels, respectively.

Full model

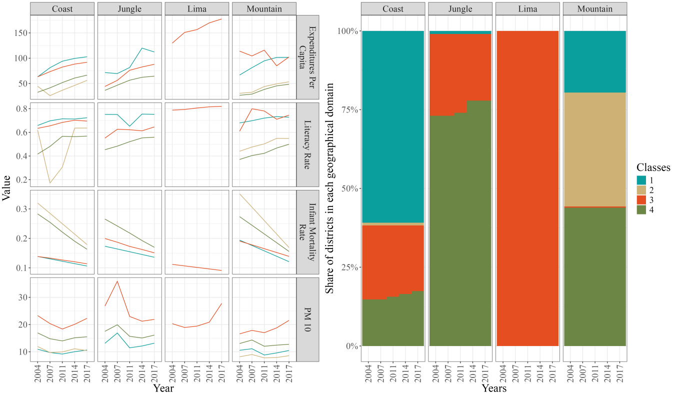


Figure: Left panel: average value of the considered variables by full model class, year and geographical domain. Right panel: the share of districts present in each full model class by geographical domain.

Conclusions

During a period of sustained economic growth in Peru between 2004 and 2017,

Conclusions

During a period of sustained economic growth in Peru between 2004 and 2017,

- The HDI model confirms the **inequality reduction** found in the literature.

Conclusions

During a period of sustained economic growth in Peru between 2004 and 2017,

- The HDI model confirms the **inequality reduction** found in the literature.
- Inequality is reduced more **from the top** than from the bottom.

Conclusions

During a period of sustained economic growth in Peru between 2004 and 2017,

- The HDI model confirms the **inequality reduction** found in the literature.
- Inequality is reduced more **from the top** than from the bottom.
- When we measure sustainable development using our Full model, **the reduction is not present antymore**.

Conclusions

During a period of sustained economic growth in Peru between 2004 and 2017,

- The HDI model confirms the **inequality reduction** found in the literature.
- Inequality is reduced more **from the top** than from the bottom.
- When we measure sustainable development using our Full model, **the reduction is not present antymore**.
- The capital city, **Lima, is not transitioning** in any specification of the model.

Conclusions

During a period of sustained economic growth in Peru between 2004 and 2017,

- The HDI model confirms the **inequality reduction** found in the literature.
- Inequality is reduced more **from the top** than from the bottom.
- When we measure sustainable development using our Full model, **the reduction is not present antymore**.
- The capital city, **Lima, is not transitioning** in any specification of the model.
- Our results suggest that if we embrace the concept of sustainable development, this implies that poverty reduction must **also** involve improvements in the environmental sphere.

Conclusions

During a period of sustained economic growth in Peru between 2004 and 2017,

- The HDI model confirms the **inequality reduction** found in the literature.
- Inequality is reduced more **from the top** than from the bottom.
- When we measure sustainable development using our Full model, **the reduction is not present antymore**.
- The capital city, **Lima, is not transitioning** in any specification of the model.
- Our results suggest that if we embrace the concept of sustainable development, this implies that poverty reduction must **also** involve improvements in the environmental sphere.
- **We do not have a causal inference scheme**; thus, our conclusions are descriptive of the observed phenomenon.

Robustness

- We compute the **subnational HDI** using the data available to us and we confront it with the results by Smits and Permanyer, 2019. We find levels with a different magnitude, but similar trends.

- The results of the full model are robust to the use of: **income per capita** instead of expenditures per capita, $PM_{2.5}$ instead of PM_{10}, NDVI instead of PM_{10}.

Thank you.

References I

Bartolucci, F., Farcomeni, A., & Pennoni, F. (2013). *Latent markov models for longitudinal data*. CRC Press.

Han, X., Guo, Q., Liu, C., Strauss, H., Yang, J., Hu, J., Wei, R., Tian, L., Kong, J., & Peters, M. (2016). Effect of the pollution control measures on pm2. 5 during the 2015 china victory day parade: Implication from water-soluble ions and sulfur isotope. *Environmental Pollution*, *218*, 230–241.

Hao, Y., & Liu, Y.-M. (2016). The influential factors of urban pm2. 5 concentrations in china: A spatial econometric analysis. *Journal of Cleaner production*, *112*, 1443–1453.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., et al. (2019). The cams reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics*, *19*(6), 3515–3556.

Ji, D., Cui, Y., Li, L., He, J., Wang, L., Zhang, H., Wang, W., Zhou, L., Maenhaut, W., Wen, T., et al. (2018). Characterization and source identification of fine particulate matter in urban beijing during the 2015 spring festival. *Science of the Total Environment*, *628*, 430–440.

Lélé, S. M. (1991). Sustainable development: A critical review. *World development*, *19*(6), 607–621.

Lou, C.-R., Liu, H.-Y., Li, Y.-F., & Li, Y.-L. (2016). Socioeconomic drivers of pm2. 5 in the accumulation phase of air pollution episodes in the yangtze river delta of china. *International journal of environmental research and public health*, *13*(10), 928.

Organization, W. H., et al. (2021). *Who global air quality guidelines: Particulate matter (pm2. 5 and pm10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization.

Porter, M. E., & Linde, C. v. d. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of economic perspectives*, *9*(4), 97–118.

San Martini, F. M., Hasenkopf, C. A., & Roberts, D. C. (2015). Statistical analysis of pm2. 5 observations from diplomatic facilities in china. *Atmospheric Environment*, *110*, 174–185.

Smits, J., & Permanyer, I. (2019). The subnational human development database. *Scientific Data*, *6*(1), 190038. https://doi.org/10.1038/sdata.2019.38

Todaro, M. P., & Smith, S. C. (2009). Economic development. harlow.

References II

Tunno, B. J., Dalton, R., Michanowicz, D. R., Shmool, J. L., Kinnee, E., Tripathy, S., Cambal, L., & Clougherty, J. E. (2016). Spatial patterning in pm2. 5 constituents under an inversion-focused sampling design across an urban area of complex terrain. *Journal of Exposure Science & Environmental Epidemiology*, *26*(4), 385–396.

Wang, S., Liu, X., Yang, X., Zou, B., & Wang, J. (2018). Spatial variations of pm2. 5 in chinese cities for the joint impacts of human activities and natural conditions: A global and local regression perspective. *Journal of cleaner production*, *203*, 143–152.

Wang, Z.-b., & Fang, C.-l. (2016). Spatial-temporal characteristics and determinants of pm2. 5 in the bohai rim urban agglomeration. *Chemosphere*, *148*, 148–162.

Wu, G., Wang, S., Jiang, Z., Hopke, P. K., Rich, D. Q., Chen, L., Lin, S., Zhang, K., Romeiko, X. X., Qu, Y., et al. (2023). Spatial-temporal pattern of tuberculosis mortality in china and its relationship with long-term pm2. 5 exposure based on a causal inference approach. *Environmental Research Letters*, *18*(8), 084006.

Appendix

Air pollution

- PM2.5 are universally considered as a **main cause for haze formation due to both natural and anthropogenic factors**. PM2.5 is closely related to human activities within the atmospheric boundary layer. PM2.5 reduces visibility and participates in chemical reactions in the atmosphere to generate new pollutants and severely affects human health.
- The results of various scholars suggest that **industrial emissions**, **car exhaust** (Ji et al., 2018), **urban dust** (Han et al., 2016), **population concentration** (Lou et al., 2016), and **biomass and coal burning** (S. Wang et al., 2018) are the main sources of PM2.5 pollution.
- Other studies have found that **meteorological elements** such as air temperature and pressure (Tunno et al., 2016), wind speed and direction (Z.-b. Wang & Fang, 2016), and precipitation and humidity (San Martini et al., 2015) are important in determining the spatial patterns of PM2.5.
- **Environmental elements** such as altitude, landscapes, and vegetation cover also have a significant impact on PM2.5 pollution (Hao & Liu, 2016).
- Additionally, numerous recent medical studies have demonstrated that P**M2.5 is highly correlated with morbidity and mortality**, and exposed individuals are at increased risk of developing lung cancer, cardiovascular disease, and other respiratory diseases (Wu et al., 2023).

back

PM_{10} and ENSO

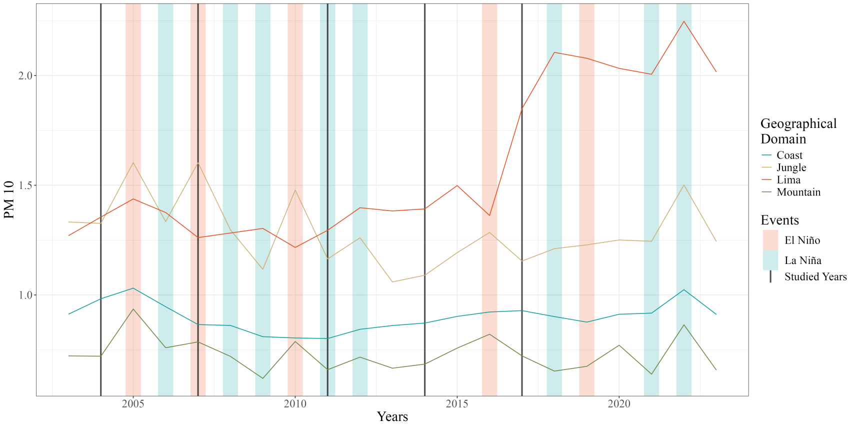


Figure: Trend in PM_{10} normalized by the WHO standard of 15 $\mu g/m^3$ by geographical domain. In red El Niño, in blue La Niña. The black line represents the years considered for our analysis.

back

Variability

Table: Descriptive statistics per year of the considered variables. The coefficient of variation (CV) is computed dividing the standard deviation (SD) by the mean value. The Median Average Deviation (MAD) considers the variables after transformation.

Year	**Mean**	**SD**	**CV**	**MAD**
	Expenditures per capita			
2004	48.35	39.21	0.81	0.72
2007	57.47	46.00	0.80	0.78
2011	68.32	43.83	0.64	0.70
2014	74.84	46.42	0.62	0.56
2017	78.22	46.20	0.59	0.55
	Literacy Rate			
2004	0.52	0.19	0.37	0.90
2007	0.55	0.19	0.34	0.85
2011	0.58	0.18	0.31	0.87
2014	0.60	0.16	0.27	0.78
2017	0.61	0.15	0.25	0.67

back

Variability

Table: Descriptive statistics per year of the considered variables. The coefficient of variation (CV) is computed dividing the standard deviation (SD) by the mean value. The Median Average Deviation (MAD) considers the variables after transformation.

Year	Mean	SD	CV	MAD
	Infant Mortality Rate			
2004	0.24	0.10	0.42	0.56
2007	0.22	0.08	0.38	0.52
2011	0.19	0.06	0.33	0.46
2014	0.17	0.05	0.28	0.37
2017	0.14	0.04	0.25	0.25
	PM_{10}			
2004	14.14	5.91	0.42	0.42
2007	14.98	7.36	0.49	0.43
2011	12.53	4.96	0.40	0.42
2014	12.84	4.72	0.37	0.37
2017	14.00	5.87	0.42	0.37

back