Internal Migration and Poverty Dynamic Analysis Using a Bayesian Approach: The Tunisian Case

Amal Jmaii, Damien Rousseliere, Besma Belhadj

Abstract—We explore the relationship between internal migration and poverty in Tunisia. We present a methodology combining potential outcomes approach with multiple imputation to highlight the effect of internal migration on poverty states. We find that probability of being poor decreases when leaving the poorest regions (the west areas) to the richer regions (greater Tunis and the east regions).

Keywords—Internal migration, Bayesian approach, poverty dynamics, Tunisia.

I. INTRODUCTION

The relationship between poverty and migration is strongly discussed in the literature [1]-[4]. Reducing poverty remain major challenges in developing countries due to the absence/scarcity of data. Some authors have tried to overcome this difficulty through a cross-sectional analysis. Reference [5] measured chronic poverty in Papua New Guinea without a panel data using a decomposed cross-sectional poverty method. References [7], [13] and many other authors determined chronic poverty on the basis of household surveys. They have used information about household living standard to distinguish between chronic and transient poverty. Several economists, such as Bourguignon et al. [5] and [11], used different methods to estimate the dynamic aspect of poverty without recourse to panel data. However, Bourguignon et al.'s study analyzed vulnerability to ex-ante poverty; it was not intended ex-post analysis of the determinants of mobility entering-exiting of poverty. In fact, this is the difference between vulnerabitlity and poverty studies. But when having two or more cross-sectional surveys, we can compare the values of poverty index calculated on the basis of each survey. In this case, we can only analyze the variation of poverty aggregate characteristics over time. Therefore, we are not able to follow poor individuals over time. The solution proposed by [9], [22]-[25] is to create panels data from averages information and taken cohorts (classified according to criteria that ensuring some homogeneity) as units. In fact, this technique allows finding some advantages of panel model. As the ability to model the dynamic effects, while avoiding the problem of "attrition", cited by [27], [14] and [15], which corresponds to the loss of individuals over time. However, the disadvantage of this method is that it can not analyse the poverty dynamics within groups, so we can not distinguish chronic poverty from transient poverty within each cohort. Perhaps because of this point pseudo-panel models have not been broadly applied to analyze poverty dynamics.

In particular, in Tunisia, the evolution of poverty shows a significant decline in national poverty from 23% in 2005 to 15% in 2010 (about 2% per year). However, poverty in rural areas remains relatively stable over the same period. Indeed, rural populations are the most affected by poverty. Furthermore, the fuzzy subset theory [6] make available a new replay to the traditional Jalan-Ravallion approach by proposing new measures for chronic, transient and persistent poverty where data are incomplete through time. They found that poverty is mainly a chronic phenomenon.

In our case where information is incomplete through time and surveys are not in form of panel data, we consider an alternative potential outcomes or counterfactual approach [19]-[21] combined with multiple imputation to get around this problem. Our proposition aims to impute potential value for non-observable variable in time t. This study presents a contribution to the extant literature, as it examines the trajectories of poverty based on the distinction between chronic and transient poverty on a rural-urban disparities context in the case of independent surveys.

The rest of the paper is organized as follows. We will begin by providing some background knowledge (Section II). We present the framework in Section III. The data source and the econometric modeling is described in Section IV. In the Section V, we will discuss the extent of poverty and regional disparity over time. And finally, the Section VI concludes.

II. THE TUNISIAN CONTEXT

As many developing countries, poverty in Tunisia is concentrated rather in rural areas and in some regions of the country, particularly the west. Households with higher level of poverty rate are more concentrated in the interior regions of the country than the inland ones. A strong variation in poverty rates between regions (Fig. 1) may be the cause of social instability and population movement. Thus, measurement of poverty at the regional level allows bettering defining the priorities for regional development. However, poverty decreased from 2005 to 2010. Figs. 1 and 2 depict the stochastic dominance curves of regions and urban-rural decomposition. We can clearly observe that the prevalence of poverty in urban area dominates rural area at every point of the distribution. Otherwise, we register in urban environment the low poverty compared to rural areas. This decrease faces...
to a higher consumption disparities with economic inequalities assert that the GDP growth was biased towards the non-poor. Until now, the adopted economic and social development does not correspond to good regional governance objectives that Tunisia should achieve. Thus, the disappointment of many Tunisians is the measure of their expectations after the revolution. They express some feeling of distrust face to the public policies which devoted little regard for social inequalities.

According to several searchers [38], [39], the precise knowledge of the extent of internal population migration is largely dependent on poverty and regional polarization levels. In Tunisia, the extent of poverty is a well-established practice. However, despite the various improvements made by the INS, the basis of this methodology is the same since 30 years ago. In fact, according to the international standards, these methods require several adjustments to take into consideration recent changes especially after the revolution. Now, following the INS reasoning, poverty level reach 32% in some regions compared to a national average level of 15.5%.

The data source used in our study is the National Survey of budget, consumption and living standard of households (NSCH) for the two years 2005 and 2010 (they are a quinquennial surveys). Samples are obtained using a stratified random sampling.

From stochastic dominance curve we can observe clearly that the prevalence of poverty in urban area dominates rural area at every point of the distribution. Otherwise, we register in urban environment the low poverty compared to rural areas. Moreover the country has recently recorded a sudden acceleration of the rate of migrants flow between regions with an annual rate of about 160 between May 2011. This rate is about 88,9 in 2004 and 50,9 in 2009. This remarkable acceleration and the convergence to the Great Tunis, leads us to pose several questions. Only Greater Tunis and Middle East noted a positive stock of migrants. According to [6], these regions are characterized by lower poverty rates. Geographical poverty trap is then a challenging task. Thus, the principal objective of this study is to identify the impact of internal migration into the probability of entry and exit from poverty.

### III. EMPIRICAL ILLUSTRATION

#### A. Data Sources and Statistic Descriptives

The empirics are based on the Tunisian household’s expenditures surveys of 2005 an 2010. To run a multiple imputation of our interest variables (expenditures employment and regions) we have selected some characteristics that are considered sufficiently stable over time namely: The sex of individuals (1 for female), the region: Urban Great Tunisia, Rural great Tunisia, urban east, rural east, urban west and rural west¹, the educational level: Illiterate (reference modality), primary level, secondary level and higher level. For age we choose four generations: Individuals who were born between 1984 and 1975, 1974 and 1965, 1964 and 1955, and finally, born between 1954 and 1945. We have not taken the household head as a reference since it may change from a year to another depending on the financial condition of the household. This method is also able to deal with internal migration problem, which was mentioned by several researchers [16]- [18].

The data source used in our study is the National Survey of budget, consumption and living standard of households (NSCH) for the two years 2005 and 2010 (they are a quinquennial surveys). Samples are obtained using a stratified random sampling. We can observe clearly that the prevalence of poverty in urban area dominates rural area at every point of the distribution. Otherwise, we register in urban environment the low poverty compared to rural areas.

Table II describes the average sample of the used variables. The statistics show that women dominate the sample with a percentage of 53.8%. Those individuals who were born between 1984 and 1975 represent 31.5% of total sample, following by those who were born between 1974 and 1965 by 27.4%, those who were born between 1964 and 1955 represent 24.9% and finally those who were born between 1954 and 1945. We have not taken the household head as a reference since it may change from a year to another depending on the financial condition of the household. This method is also able to deal with internal migration problem, which was mentioned by several researchers [16]- [18].

Table II describes the average sample of the used variables. The statistics show that women dominate the sample with a percentage of 53.8%. Those individuals who were born between 1984 and 1975 represent 31.5% of total sample, following by those who were born between 1974 and 1965 by 27.4%, those who were born between 1964 and 1955 represent 24.9% and finally those who were born between 1954 and 1945 represent 16.1%. Individuals who are illiterate represent 21% of total sample and those who have a primary educational level represent 38.6%. Individuals who have secondary educational level represent 29.4% and those who have a higher educational level represent 10.7%.

¹We choose this decomposition of areas since we are interested by the urban rural disparities over time. Moreover, we could use any other figure.
We consider an individual as poor if the equivalent expenditure of household is below the fixed poverty line2 (we define the poverty line as the 50% of the mean distribution). We fix two poverty lines for each year. Evidently, we use the CPI (the Consumer Prices Index)3 to compare expenditure of households in dinars, over time.

In this paper total expenditure is taken as a standard of living indicator. Indeed, consumption expenditure is further characterized by their stability over time compared to income fluctuations. They provide information about the degree of satisfaction that comes from the consumption of goods and services. This approach has been advocated in recent studies by [28], [26].

We consider an individual as poor if the equivalent expenditure of household is below the fixed poverty line2 (we define the poverty line as the 50% of the mean distribution). We fix two poverty lines for each year. Evidently, we use the CPI (the Consumer Prices Index)3 to compare expenditure households in dinars, over time.

### B. Estimation

Potential outcome framework demonstrates that we can treated causal inference as a particular case of missing data problem [32], [8]. Following [33], we suppose $M$ to be the indicator matrix of missing data and $\gamma = (\mu, \Sigma)$ the parameters of the model. We assume that the data are in MAR4 (Missing at random), and we use the amelia II algorithm to impute the missing variable. This algorithm combines the classic algorithm (EM)5 with the bootstrap approach, we assume that:

$$ p(D^{\text{ob}}, M|\gamma) = p(M|D^{\text{ob}})p(D^{\text{ob}}|\gamma) $$ \hspace{1cm} (1)

We write the likelihood as:

$$ L(\gamma|D_{ob}) \propto p(D_{ob}|\gamma), $$ \hspace{1cm} (2)

and

$$ p(D^{\text{ob}}|\gamma) = \int p(D|\gamma)dD_{mis} $$ \hspace{1cm} (3)

The posteriori law is then defined

$$ p(\gamma|D_{ob}) \propto p(D_{ob}|\gamma) = \int (D|\gamma)dD_{mis} $$ \hspace{1cm} (4)

For each draw, the data are estimated by bootstrap in order to simulate the uncertainty, and then the EM algorithm is executed to find the posterior estimate $\hat{\gamma}_{MAP}$ for bootstrapping data [34], [35]. The study uses a recursive bivariate probit model [12]. The main reasons to use this model were twofold: On the one side, it verifies if poverty causes future poverty through the introduction of “poverty 2005” variable in the first equation, and on the other side, it detects unobservable effects that can be analyzed with the sign and the significance of the autocorrelation term ($\rho$). In this section, we use a dynamic recursive bivariate probit to estimate poverty dynamics. The

---

2We use per capita household expenditure and we consider that all members of the households have the same weight in the denominator. Then, we divide total household expenditure by the square root of number of members (in this case, square root was used as scaling factor)

3Since the CPI measures changes in the general price level, it is often used to estimate changes in purchasing power of Tunisian dinars.

4In the case of MAR data, the probability of absence is related to one or more other observed variables ([19])

5The classic algorithm of maximum of likelihood [36]
model is able to take into consideration the past experience of poverty. Following the specification given by [12] and [10], we assume that:

\[ P_1^* = \theta_1 + v_1 \]
\[ P_2^* = \theta_2 + v_2 \]
\[ \theta_1 = \beta_1 x_1 \]
\[ \theta_2 = \gamma_1 + \beta_2 x_2 \]
\[ p = g(p^*) = (1(P_1^* > 0), 1(P_2^* > 0))' \]
\[ \epsilon = (v_1, v_2)' \sim N(0, \Sigma) \]

(5)

Note that \( g(.) \) is the link vector-valued function where \( \beta_1 \) and \( \theta_1 \) represent matrix of coefficients \((i=1,2)\) \( P_1^* \) and \( P_2^* \) are consecutively poverty of 2005 and poverty of 2010 and \((v_1, v_2)\) the residuals vector. \( I \) represents the indicator function, \( \rho \) measures endogeneity of \( p_1 \) in the \( P_2^* \) equation. In addition, the two equations of the model follow the bivariate normal distribution that is assumed by the following joint density function:

\[ \phi_2(v_1, v_2, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp[-\frac{1}{2(1-\rho^2)}(v_1^2 + v_2^2 - 2\rho v_1 v_2)] \]

where \( E(v_1, x_1, x_2) = \lambda_1 = 0 \) and \( E(v_2, x_1, x_2) = \lambda_2 = 0 \)
\[ \text{var}(v_1, x_1, x_2) = \sigma_1^2 = 1, \text{var}(v_2, x_1, x_2) = \sigma_2^2 = 1 \]
\[ \text{cov}(v_1, v_2, x_1, x_2) = \rho \text{ with } -1 \leq \rho \leq 1 \]

The likelihood function of the model is based on four parts:

\[ L(\beta_1, \beta_2 \mid x_1, x_2) = \prod_{p1} P_{P10} P_{P01} (1-P_{P10}) P_{P00} (1-P_{P01}) (1-P_{P1}) (1-P_{P0}) \]

where:

\[ p_{11} = P_r(P_2 = 1, P_1 = 1 \mid x_1, x_2) = \phi(x_1 \beta_1 + \gamma, x_2 \beta_2, \rho) \]
\[ p_{10} = P_r(P_2 = 1, P_1 = 0 \mid x_1, x_2) = \phi(x_1 \beta_1 - x_2 \beta_2, -\rho) \]
\[ p_{01} = P_r(P_2 = 0, P_1 = 1 \mid x_1, x_2) = \phi(-x_1 \beta_1 - \gamma, x_2 \beta_2, -\rho) \]
\[ p_{00} = P_r(P_2 = 0, P_1 = 0 \mid x_1, x_2) = \phi(x_1 \beta_1 + \gamma, x_2 \beta_2, \rho) \]

IV. RESULTS

Results show that individual how poor in 2005 has a higher probability to be poor in 2010 (since the auto-correlation coefficient \( \rho \) is statistically significant). With regard to the parameter linked to poverty status of 2005, the marginal effect indicates that the probability of being poor at 2010 depends on the probability of being poor at 2005. This is meaning that a poor person at 2005 have a higher probability of being poor at 2010 regarding to non-poor individuals.

With regard to the generation variables, we found that all coefficients are statistically significant and negative for the older generation. This sign indicates that the probability of being poor is greater for younger individuals. This individuals are less likely to be initially poor than the other generations. This find might be explained by the fact that this generation was more formed in terms of education system compared to others. As a result they had more chance of having an employment than others. As expected, we find that the parameter of initial condition (poverty 2005) is significantly positive, the marginal effect of this variable indicates that the probability of being poor is higher when the individual has experienced poverty in the past.

Finally, our study highlights that 61% (0.3/0.49) of total observed poverty (0.49) is the origin of a stationary propension of poverty (i.e. chronic condition) while 39% coming from transient poverty (0.19/0.49). When we compare different poverty lines, we highlight similar results. As expected, our results highlight a highest observed poverty as well as transient poverty and chronic poverty in the rural west region. This is logical since these regions suffer from a bad social conditions and a wicked infrastructure that prevents them to escape from poverty. Regional disparities in the Tunisian labor market, have convinced many potential job seekers to change the place of residence to the great Tunis or to the urban east.

Regarding to the educational level effects, we point the similar probabilities for both chronic and transient poverty for individuals whose educational level is primary or illiterate (40% of the observed poverty). This groups might suffer from periodical changes in their status. For the variable sexes of individuals, we underline similar probability results with a little difference for women. In fact, men present higher chronic level of poverty faced to higher transient poverty for women. We can explain this by the fact that on labor market, men are facing more difficulties to find a decent job than women. In fact, the rate of women who have a high level of education overcomes the men’s rate, while the rate of non-schooling of men exceeds that of women. These two factors may decrease the chance of getting a permanent job.

To analyze the effect of region of residence on the persistence of poverty and the downward into poverty, we propose to predict some other simulated cases based on the estimated coefficients of our model (Table II). Unemployment is considered as a condition related to the concept of vulnerability to poverty rather than to a chronic poverty state. This result is justified when simulations (1) and (6) are compared. Such finds are expected as unemployment is theoretically approved as a transient state conditioned by the short-term situation of the national economy.

Regarding Table II, the probability of transient and chronic poverty of an individual who initially and originally lived in rural west areas decreases when he changes residence to the urban Great Tunis. When we change the educational level to secondary level for the same individuals, their probability decreases compared to basic case. Great Tunis represents the destination of all individuals looking for a decent job. In
Table I presents more detailed information about poverty transition between regions. In Tunisia, significant sending regions as a share of their 1995 population included the Urban East, Urban West and Central regions. Upper West migrants tended to go to the Brong Ahafo and Ashanti regions primarily, whilst Upper East migrants went to the Ashanti, Brong Ahafo and Western regions. This is strongly suggestive of migration from less economically successful to more economically successful regions. However, it is also evident that geographical distances along with other drivers shape patterns of mobility in Ghana, since migrants from the poorer Upper West and Upper East regions more often go to the relatively nearby regions of Brong Ahafo and Ashanti, rather than to the richer but more distant regions of the coastal belt. This is also consistent with previous evidence (see [31]). Similarly, more educated and wealthy people from the Upper West prefer to migrate to urban centres including Accra while the poor and illiterate migrate to the Brong Ahafo region. The range of economic opportunities available in some of the regions is also another important factor determining the choice of destination [30].

Table III presents model results. Explanatory variables include sex, education level, employment status, region, generation, and poverty status. Marginal effects and standard errors are provided for both chronic and transient poverty. The table shows that male sex, primary education, salaried employment, rural Great Tunisia, second generation, and poverty in 2005 are significant predictors of chronic poverty. On the other hand, salaried employment, rural Great Tunisia, and poverty in 2005 are significant predictors of transient poverty. The T-statistic is 22.14 with 151 degrees of freedom, indicating statistical significance.

V. CONCLUSION

In this study, we have advanced a Bayesian measure by combining potential outcomes approach with multiple imputation. Most works on poverty dynamics, in developing country, consisted in measuring and identifying correlated variables to chronic and transient poverty. Further research is necessary to understand what causes chronic poverty and why some groups of individuals are not able to accumulate assets that generate sufficient income to increase consumption expenditures above a minimum acceptable level (i.e. poverty line). To execute such analysis panel, data are recommended. However, as in most developing countries, penalized data about individual’s well-being conditions do not exist. As a result, researchers are able to conduct a ex-post dynamic analysis of poverty. This paper proposes to use potential outcomes approach as a particular case of missing data to impute potential variables and, in a second step, to run a recursive biprobit model. The main objectives of our study were, first of all, the analysis of poverty dynamics by analyzing its chronic and transitory characteristics in order to improve knowledge about this phenomenon. Secondly, we identify more efficient policy measurements in favor of population suffering from persistent poverty. We have applied this approach to analyze internal migration and poverty in the case of Tunisian households from 2005 to 2010. In sum, potential outcomes and causal inference appear to be successful methods to run a dynamic analysis of poverty. Our
finding emphasizes, in contrast to the INS’s results, that urban
Tunis is not the less vulnerable to poverty region. It is
rather the urban west. This find is justified by the effect
of internal migration. Indeed, a higher percentage of poor
individuals who live in rural areas change their residence
to go to the great Tunis where they hope to find a job
opportunities and improve their well-being. But over time,
these immigrants will constitute the vulnerable part of this
region. The importance of our analysis, compared to others
works related to Tunisian context, comes from the fact that
it highlights the internal migration movement. This variable is
interesting in the study of poverty dynamics because it changes
the traditional figure of the Tunisian regional poverty.

This paper shows interesting results in terms of public
policy. Firstly, the recorded increase of internal migration
refers to the need of rebalancing regional development in the
country. On the other hand, since the presented methodology
highlight that Tunisian poverty is mainly chronic, an effective
policy to reduce poverty must involve a multi-sectorial
reform with a constructively program of income inequality
distribution and reduce regional disparities are required. Such
programs are able to improve infrastructure in rural areas,
facilitate access to information, labor market and public
services. Potential extensions of this methodology may include
larger database about internal migration within the framework
of a deeping vision of internal migration and poverty concept.

REFERENCES

[1] B. Barham and S. Boucher, Migration, remittances, and inequality:
estimating the net effects of migration on income distribution Journal
of migration and inequality: theory and evidence from Mexico Journal
of migration: Evidence from Poland Economics Letters: Elsevier, 2009,
pp. 119-122.
migration and interregional convergence in Russia Journal of comparative
Evidence From Tunisia International Journal of Scientific and Technology
[8] G. W. Imbens and D. Rubin, Causal inference in statistics, and in the
social and biomedical sciences Cambridge University Press New York,
2009.
[9] F. Gardes and S. Langlois and D. Richaudieu, Cross-section versus
time-series income elasticities of Canadian consumption Economics
[10] D. Roodman, An expanded framework for mixed process modeling in
Stata German Stata Users’ Group Meetings: Stata Users Group, 2013.
before and after the Economic Crisis Asian Economic Journal: Wiley Online
Library, 2003, pp. 45-64.
[12] D. Roodman, Fitting fully observed recursive mixed-process models with
cmp The Stata Journal, 2011, pp. 159-206.
Centre, 2002.
adjustment in later life: evidence from the British Household Panel Survey
[16] H. De Haas, The internal dynamics of migration processes: a theoretical
journal of ethnic and migration studies: Taylor & Francis, 2010,
pp. 1587-1617.
[17] M. A. Collinson, Striving against adversity: the dynamics of migration,
health and poverty in rural South Africa Global Health Action:Co-Action
[18] P. Deshgankar, Internal migration, poverty and development in Asia:
Including the excluded IDS Bulletin: Wiley Online Library, 2006,
pp. 88-100.
augmentation: Comment: A noniterative sampling/importance resampling
alternative to the data augmentation algorithm for creating a few
imputations when fractions of missing information are modest: The SIR
algorithm Journal of the American Statistical Association: JSTOR, 1987,
pp. 543-546.
[20] D. Rubin, Estimating causal effects of treatments in randomized and
nonrandomized studies. Journal of educational Psychology: American
experiments: Essay on principles Statistical Science: American
Psychological Association, 1923, pp. 465-480.
[22] A. Deaton, Panel data from time series of cross-sections Journal of
[23] M. Browning and A. Deaton and M. Irish, A profitable approach to labor
supply and commodity demands over the life-cycle Journal of econometric
subjective et indexée sur leur revenu Economie et statistique: Institut
durables and liquidation constraints: A cohort analysis European Economic
[26] R. Chamarbagwala, Economic liberalization and urban–rural inequality
in India: a quantile regression analysis Empirical Economics: Springer,
[27] Z. Bahlali and Q. Li, Testing AR (1) against MA (1) disturbances in
133-151.
Inequality in Thailand: 1990-2006 Asian Economic Journal: Wiley Online
[29] H. Hertelli, Émigration et immigration des Tunisiens en Europe.
Migrations societe: Centre d’information et d’études sur les migrations
[30] A. McKay and P. Quartey, Evolution of spatial inequality between
northern and southern Ghana mimeo, 2008.
[31] C. Ackah and D. Medvedev, Internal Migration in Ghana: Determinants
Using Semi-Parametric Imputation Proceedings of the Section on Survey
1085-1096.
Judkins-Westa, Causal Inference Using Semi-Parametric Imputation Proceedings
of the Section on Survey Research Methods of the American Statistical Association,
[34] A. Blackwell and J. Honaker and G. King, Amelia II: A program for
[35] A. Blackwell and J. Honaker and G. King, A Unified Approach to
Measurement Error and Missing Data: Details and Extensions Sociological Methods and Research, 2015.
from incomplete data via the EM algorithm Proceedings of the EM algorithm
[37] J. P. Payet, Les ´ecoles en milieu d ´efavoris´e dans le Grand Tunis. Un
[38] J. A. J. Van der Klaauw, The analysis of household surveys: a microeconometric