Does corruption influence young brain drain? Evidence from Italy.

Alessandra Patti[1], Emanuele Millemaci[2]

1. Ph. D Student of Economics, Department of Economics, University of Messina
2. Associate Professor of Economics, Department of Economics, University of Messina

Abstract

In the last couple of decades, young skilled flow within Italian provinces has begun to increase at higher speed than ever. While it is probably too early to say if this process is transitory or permanent, it is undoubted that it is important and needs to be deeply studied by researchers as well as constantly monitored by policy makers. Previous empirical studies have demonstrated that skilled migration is influenced by economic factors, such as income per capita and employment of origin and destination places, and, with a less extent, by the search of places endowed with more amenities. In the crossroad between these two factors, this paper investigates the role of corruption as a push factor explaining skilled flows. The empirical analysis mainly relies on the estimation of a gravity equation by using both the methods of Zero-Inflated Poisson and Pseudo-Poisson Maximum Likelihood with High Dimensional Fixed Effects with bilateral data of the Italian universities and provinces where they are located. Results suggest the existence of both push and pull mechanisms at play, as high corruption incentivizes skilled emigration toward destination provinces that, instead, exhibit lower levels of corruption. Moreover, sensitivity of the prospective tertiary students towards corruption varies depending on the fields of study of interest. Finally, we consider the effect of corruption over skilled flows from the South tod the Centre-North, partly motivating the existence of longstanding socio-economic differences between the North and the South of Italy.

Keywords **brain drain**, **corruption**, **panel data**, **zip**, **ppmlhdfe**, **gravity**

JEL Classification: **D73 F12 R23**

I. **Introduction**

Corruption represents an economic loss that harms, inter alia, the wealth of an economy. Thus, the derived amount of waste represents economic and social failures to deal with.

Recent empirical research has demonstrated that high corruption occurrence negatively affects the level of growth (Lisciandra and Millemaci, 2016), employability (Poprawe, 2015), and overall quality of local government (Nifo and Vecchione, 2014; Charron et al., 2013), for any given country. On the other hand, countries with lower levels of corruption are characterized by higher dynamism of labour market that facilitates more equitable economic mobility within participants (Dotti et al., 2013; Mayda, 2009). Besides, dishonest bureaucracies slow down the speed of doing business and burden skilled individuals with monetary and non-monetary costs. Thus, corruption forces part of the labour force out of the corrupted country, thereby lowering output even further (Poprawe, 2015; Dotti et al., 2013).

On the other hand, apart from the key role played by economic and occupational motivations, recent research has witnessed the increasing interest toward the effect of social factors, such as democracy, trust in legal authorities, awards for human capabilities and other cultural-dependent amenities, that encourage skilled migration to move away from origin areas (Auer et al., 2018; Beine et al., 2013). All aforesaid studies are performed on cross-regionals and panel data basis for OECD countries and for Italian regions. Hence, young skilled migration between southern and centre-northern provinces of Italy is an up-to-dated phenomenon that has inspired our research because it deserves particular attention and it has many economic and socio-cultural implications.

From 2010 to 2017, the phenomenon of Italian skilled migration within provinces revealed a positive trend. In fact, the interprovincial skilled emigration rate, conceived as the ratio of the total enrolled skilled students who move to other destinations, over the total number of enrolled skilled young people who migrate and do not migrate, registered an abnormal increase of 4%. Thus, studying why this phenomenon changes speedily in relative low time became a socio-economic task of national importance to deal with.

From macroeconomic viewpoint, it is important to understand the dimension of the Italian skilled migration respect to the phenomenon that occurs in other countries in order to assess its effects on income, growth rates, public expenditure and fiscal revenues. From socio-economic view, the existence of the well-known multifaced dualism between centre-northern and southern places of Italy raises a lively consideration on the effects of skilled migration in terms of losses for southern provinces. In fact, further southern skilled flows may have a detrimental effect on the process of convergence, by deepening socio-economic disparities and creating negative externalities. Besides, the consistent presence of organized crime and corruption in the South may foster this inequality condition and render troubling the process of recovery.

Also, this dualism may be analysed from educational viewpoint. In fact, we are interested in evaluating the effect of young out-migration on human capital in the South, because enrolment in tertiary education in Southern area is depressed by higher pre-graduation migration because a share of skilled people prefers to pursue their studies in the destinations provinces that offer not only good didactic programs but also advantageous job opportunities. In fact, as Ciriaci (2013) reported, most part of the students, who move to destination provinces, have the likely possibility to remain to the place, where they graduated, for job reasons.

Finally, few papers have recently recognized the role of corruption, sometimes distinguishing the relative importance of its push and the pull effects.

The aim of this paper is to offer a relevant contribution to the existing literature that uses traditional Gravity models to examine skilled emigration fuelled by local macroeconomic features, such as GDP per capita, employment, governmental quality, network effects, job vacancies and housing prices (Beine et al, 2014; Dotti et al., 2013; Van Bouwel and Veugelers, 2013; Mayda, 2009), by considering corruption as main factor of influence for skilled migration from origin to destination provinces.

For this purpose, our research uses panel data aggregated at Italian provincial level and implements count data methods that relies on Poisson and Pseudo-Poisson Maximum Likelihood models within gravity set-up. The decision to adopt a gravity model is not novel in the literature on trade and migration. Research conducted by Beine et al., (2014) Van Bouwel and Veugelers (2013), Rodriguez et al., (2010) and Mayda (2009) have already used this approach to estimate international student migration's determinants across OECDs in last thirty years. Thus, this research makes a three-fold contribution on examining the relation between corruption and young skilled flows within Italian provinces in the last decade.

First, we analyse whether corruption incentives young skilled individuals, when they face the decision to enrol in university, to move to other provinces that exhibit lower levels of corruption and overall better quality of life. We presume that this happens because young skilled individuals, at the beginning of their university career, are more sensitive to fairness and legal issues for living better-off in society. Besides, young people know that their opportunities of career depend upon the competitiveness and meritocracy of their local labour market; in fact, they are looking forward to having chances to live in places where equal rewards for their capabilities are likely to occur. The idea behind is that corruption and sensibility to this phenomenon, as well as quality of university and living standards, is meaningful because it contributes to shape markedly the students' main reasons to move or rest once they decide to enrol to university.

Second, this study investigates in an indirect way to what extent investments in education in corrupted origin province might be relatively futile, unless the corruption is simultaneously dealt with, and could benefit other provinces when young skilled people decide to emigrate to universities of other destination places. Thus, investments in education in source province riddled with corruption might fade away and become a positive externality for destination provinces (Beine et al., 2014).

Finally, this research offers recommendations to policymakers to promote political interventions that improve transparency and social cohesion needed for attracting and/or retaining young skilled individuals in their own provinces. Fighting against corruption and young brain drain is not only an important short and medium-term policy concern but has even more relevant implications in the long run, as it might have long-lasting effects not only on human capital accumulation but also on economic growth. In fact, young brain drain caused by corruption can be considered urgent governmental issues to be re-addressed for improving the economic welfare of southern Italian provinces that are affected by these two phenomena mostly.

The work is structured as follows. In Section 2, we refer to the core literature that represents the basis of our research hypothesis. Section 3 describes data and methodology. Section 4 discusses the results while Section 5 concludes.

II. **Literature Review**

This section develops the conceptual foundation of our paper. In the process, it derives main research hypotheses of interest. We connect our study to recent literature on gravity models on macroeconomic factors that determine skilled migration across OECDs and Italian regions. Different empirical research has demonstrated that the increase of corruption within institutions encourages skilled emigration (Poprawe; 2015; Nifo and Vecchione, 2014;) to destination places that exhibit lower levels of corruption and overall better quality of life (Poprawe, 2015; Beine et al., 2014; Dotti et al., 2013; van Bouwel and Veugelers, 2013; Mayda, 2009). In fact, corrupted environments promote skilled emigration and discourage skilled immigration because they provide uncertain economic conditions and worse standards of living (Poprawe, 2015). Young skilled individuals move in search of higher quality of life, meant to be the overall product of a mix of economic and socio-cultural factors related to welfare. Thus, social mobility, efficiency of institutions, effectiveness of public administration and widespread civic sense are important elements that encourage skilled people to leave their origin places where aforementioned factors are relatively low (Nifo and Vecchione, 2014).

However, scholars provide attention to other socio-cultural factors that play crucial roles in skilled emigration as well. Students make migration choices according to the dynamism of the local labour market. For this reason, Dotti et al., (2013) assert that skilled emigration happens mainly due to the prospects of job vacancies for graduates rather than evaluating local quality of university where students intend to enrol. Opposite to Dotti et al. (2013), van Bouwel and Veugelers (2013) demonstrate that quality of education and better living conditions of source and destination countries significantly shapes the magnitude and direction of student mobility across regions. Higher quality of university and quality of life of source provinces redeem young brain drain to destination countries and vice versa. In addition, Mayda (2009) demonstrates that brain drain is likely to occur when there is strong asymmetry of pull and push factors between origin and destination regions.

Several empirical studies have demonstrated that bilateral skilled migration is influenced by socio-economic factors, such quality of institutions of origin and destination places. Cross-country studies have exploited quality of government, that encompasses voice and accountability, rule of law and governmental effectiveness (Nifo et al., 2013) as main variable that induces skilled flows from origin places. However, in our study, the quality of local government is captured by corruption, considered as the ideal leverage variable for analysing young Italian brain drain within provinces. Corruption may hamper the overall well-functioning of local contexts because it hinders talented and clean activities and supports public projects that are less labour intensive and more capital intensive (Corrado et al., 2018). People with higher levels of educational attainments perceive corruption as impediment to develop their own career successfully and decide to move to lands where merit is recognized fully (Bonanno et al., 2020). In addition, the perception of corruption varies according to the level of skills of individuals (Cooray and Schneider, 2016). However, there is still no evidence if on-going academic background chosen by young students reflects their sensibility toward corruption and their decision to move away. Hence, we want to proof if:

H_1. Higher level of corruption behaves as main push factor of brain drain from origin provinces, while lower level of corruption behaves as pull factor of brain gain at destination provinces

$H_{2:}$ Sensibility to corruption varies within young Italian students who enrol to academic courses belonging to different fields of study (Social Science, Physical Science and Life Science)

H3: Corruption contributes to determine long-distance skilled migration within Italian provinces

III. **Data and Methods**

Italian brain drain was studied by constructing an original panel dataset that is balanced and consists of 20,808 total observations with university as well as aggregated data at provincial level made of the combination of up to the 57 Italian provinces (N) where at least a university is present (if more than a university is present an overall measure is taken for university variables) of the origin (i) and the same 57 destinations (j), for the eight years in the time interval 2010-2017 (T). In our main analysis, we only focus on bilateral data where there is a strictly positive number of students migrating from the origin province to the destination province and where the origin province is endowed of at least a university. The reason that motivates the adoption of restricted sample of provinces with local universities relies on examining the voluntary young brain drain not forced by the absence of local university but by other non-specified reasons, on which we are interested, differently from the young brain drain that affects provinces that have not local universities and skilled emigration becomes justifiable. We do also exclude from the analysis those students that move away from Italy to attend tertiary education abroad, which would be beyond the focus of the study and not adaptable using the same empirical strategy employed for studying the within country students' movements.

The main data source used for brain flow, which is denoted by count dependent variable, is the Italian Education, University and Research Department (MIUR), while data on provincial corruption cases made by known and unknown authors, data on quality of Italian universities and data on provincial macroeconomic features are taken from different sources such as the Italian Institute of Statistics with joint union of Italian Justice Ministry (RE.GE Istat), the Italian Centre for Investments and Social Studies (Censis), the Italian University Group of ALMALAUREA and Italian Institute of Statistics (Istat) respectively. A description of each variable with its own source is provided in Table 1 below:

***Table 1**. Descriptive Statistics*

Variables	***Notes***	***Source***
enrolled	number of resident students who enrol from origin province, with one local university, to university of destination provinces	MIUR
enrol erc1	number of resident students who enrol to courses of Social Science	MIUR
enrol erc2	number of resident students who enrol to courses of Physical Science	MIUR
enrol erc3	number of resident students who enrol to courses of Life Science	MIUR
time	time expressed in minutes to travel by car from origin to destination	ISTAT
pop/pop_j	average annual population of origin/destination	ISTAT
corr_tot/corr_tot_j	corruption of origin against PA(artt. 314-322 Italian tort law)	RE.GE ISTAT
rgdppc/rgdppc_j	real GDP per capita per origin 2010-2017, base GDP year 2010	ISTAT
employment/employment_j	employment rate of origin/destination	ISTAT
uni size/uni_size_j	number of enrolled =1 small, =2 medium, =3 large origin/destination	CENSIS
zquniv/zquniv_j	standardized value of quality of university of origin/destination	AlmaLaurea
zqlife/zqlife_j	standardized value of quality of life for origin/destination	ISTAT
Dairport/Dairport_j	dummy for airport, D=1 airport presence, D=0 otherwise	ISTAT
DTAV/DTAV_j	dummy for high-speed train, D=1 HST presence, D=0 otherwise	Ferrovie Stato
Dport/Dport_j	dummy for port, D=1 port presence, D=0 otherwise	ISTAT
Dnorth/Dnorth_j	dummy for North macro area origin/destination	ISTAT
Dcentre/Dcentre_j	dummy for Centre macro area origin/destination	ISTAT
Dsouth/Dsouth_j	dummy for South macro area origin/destination	ISTAT
disccrim/disccrim_j	variable indicator of law enforcement of origin/destination	ISTAT
probofconv/probofconv_j	variable indicator of law enforcement of origin/destination	ISTAT

To estimate the brain movements, we use non-negative count dependent variable that indicates the number of skilled students who decide to enrol to their local university and/or to universities located to different destination provinces (j). In addition, we insert non-negative count-dependent variables, denoted as ERC-1, ERC-2 and ERC-3, that indicate the number of enrolled students according to the study-area chosen at university: ERC-1 stands for courses belonging to Social Sciences, ERC-2 for Physical Sciences and ERC-3 for Life Sciences.

Then, independent variables inserted are two-fold featured for origin and destination provinces, respectively. The main independent variable is the total number of corruption cases (corruption total), the sum of cases with known and unknown authors, made against public administration (P.A) according to artt. 314, 316. -bis, -ter, 317, 318. 319, -bis, -ter, -quarter, 320, 321, 322, bis of Italian tort law.

Based on recent studies conducted on skilled migration with Gravity specification (Beine et al., 2014; Dotti et al., 2013; Mayda, 2009), we insert mass variables and distance indicators in our model too. In fact, we use the variable *time*, expressed in minutes, that indicates time necessary to travel the distance occurring from source to destination provinces and the average annual population rate that indicates the magnitude and attractiveness of origin and destination provinces. Then, real GDP per capita per provinces 2010-2017, with base GDP for year 2010, is added as valid control for scale and wealth-effects. Also, employment rate is inserted because can be considered as potential factor of generating and/or attracting skilled flows within Italian provinces.

Furthermore, indicators of quality of university are included in our study. First, we use size of university, which is a categorical variable that indicates how much large is an university ; in fact, small-sized university presents lower availability of infrastructures (such as libraries, study-rooms or cafeteria) as well as didactic courses with consequent expected lower attractiveness for students. Then, we insert a new variable for quality of university, which is obtained by own elaboration using ALMALAUREA data of Italian universities. We derive the average of standardized quality of university by averaging the standardized values of age, grade, expected income per capita, expected time to find a job and employability when students reach their academic degree. All values are considered source of quality of university because are adaptive to current Italian students' preferences and expectations to enrol to one university rather than another one and to find a job according not also to their capabilities but also to the fairness and dynamism of local labour market supply. Then, we add novel variable for quality of life, constructed by own elaboration and by taking into consideration seven out of the eleven caterpillars that are commonly used to build-up the OECD Better Life Index (Income and Wealth, Job and Earnings, Housing Conditions, Health, Work-Life Balance, Education, Environmental Quality). In fact, we develop the average standardized quality of life variable by averaging the standardized values of mortality rate, working formation, gender difference in employment, the presence of green urban areas and the presence of essential public services related childhood and healthcare within Italian provinces. Undoubtedly, provinces that present better standards of living conditions are more enticing than provinces that show off lower levels of it. Moreover, we add variables of law enforcement that are relevant for skilled emigration purpose and assume double significance as crime control measurements. First, efficient judicial system prevents illegal behaviour and guarantees safety of individuals and society by enforcing controls and convicting authors of crimes. Furthermore, the efficiency of law enforcement can contribute to discover the underestimated number of potential but not already detected and convicted crimes. Thus, law enforcement becomes more precise measurement variable for controlling both effective and potential authors and criminal cases.

In addition, dummies for infrastructures features (airport, port, and high-speed railway presence) are considered because their presence constitute incentives of mobility between provinces. While the time variable only provides an estimate of by car travel distance from source to destination provinces, dummies for presence of infrastructures, exploring for the presence of alternative means of transport with respect to the highway one, allows us to control in a more accurate way for the actual time distance. Hence, a highly connected province is more attractive than the one which has low transport connections and is in not easily accessible region.

Finally, fixed effects of macro areas for North, Centre and South of Italy are inserted as well as years.

Theoretical Framework

While the interest of literature on bilateral trade flows has often focused on economic factors such as income per capita and job opportunities of origin and destination places, as proved by cross-section (Dotti et al., 2013; Van Bouwel and Veugelers, 2013) and panel data (Beine et al., 2014; Mayda, 2009) studies of periods 1995-2007 and 1985-2007 respectively, by contrast, we evaluate the influence of corruption cases on Italian brain drain.

In this manner, we can better assess the role played by corruption on Italian skilled flow and presume why southern-skilled diaspora toward northern provinces occurs usually, and, whether corruption could have higher impact on resident skilled individuals, when they prefer to enrol to universities located in different provinces. In addition, understanding this phenomenon permits us to identify an appropriate policy intervention to mitigate skilled emigration from southern provinces and build-up an equal skilled distribution within areas that would has beneficial economic implications. Moreover, corruption may favours not only unbalanced but also balanced equilibrium of brain flow. In fact, southern provinces, with higher levels of corruption, hold people who are low-sensitive to illegal activities, while northern provinces, with lower levels of corruption, embrace people who are high-sensitive to criminal facts, increasing the failure of equally distributed fair contexts. However, this unbalanced brain flow is not inexorable. In fact, it may converge because for people who leave their origin places, there are other ones who remain and are encouraged to pursue better quality of living, improving the efficiency of their societies at glance.

Hence, students' migration within Italian provinces can be considered as a form of spatial interaction. For this reason, it is plausible to study this phenomenon with the framework of gravity models. In analogy with Newton's law of gravity, resident student flows can be predicted according to the following formula:

$$I_{ij} = K \frac{M_i^{\beta_1} M_j^{\beta_2}}{d_{ij}^{\beta_3}} \qquad [1]$$

Where Iij represents the interaction intensity or the number of resident students of origin province i enrolling to university located in province j also, K is a proportionality constant, Mi is the mass of the province of origin, Mj is the mass of the province of destination, dij is the physical distance between the two provinces. In addition, β1 is the potential to generate flows from origin province (*i*), β2 is the potential to attract flows from destination province (*j*) and β3 is an impedance factor reflecting distance decay that can be figured out by lower mobility connections due to inefficient transport infrastructures.

Empirical Model

Given Equation (1), we estimate an econometric model with the following form:

$$Enrolled_{ij} = f\,(\text{time } pop_{ij}\ corruption_{ij}\ rgdppc_{ij}\ employment_{ij}\ unisize_{ij}\ quniversity_{ij}\ qlife_{ij}\ infrastructures_{ij}\ law_{ij}) \quad [2]$$

Hence, Equation (2) states that young students' emigration will be made upon evaluation on current macroeconomics factors of both origin and destination provinces.

For this purpose, our study empirically present Equation (2) by adopting Zero Inflated Poisson (ZIP) and Pseudo-Poisson Maximum Likelihood (PPML) with multiple High Dimensional Fixed Effects (HDFE) models used by Long (1997), Correia et al. (2020) and Correia (2016), respectively.

Zero-Inflated Poisson allows us to takes into consideration the highly skewed distribution of our dependent variable and allows for overdispersion assuming that there are two different types of observations in the data: i) those who have a zero count with a probability of 1 (0 group), and ii) those who have counts predicted by the standard Poisson (not 0 group). Observed zero could be from either group and if the zero is from the 0 group, it indicates that the observation is free from the probability of having a positive outcome (Long, 1997).

In addition, we handle Pseudo-Poisson Maximum Likelihood with multiple High Dimensional Fixed Effects (PPMLHDFE) and then, we compare the results obtained by with kwon Zero-Inflated Poisson model. Presenting both these models appears useful as there are also several advantages related to the method of the Pseudo-Poisson Maximum Likelihood as: i) relaxing assumption of knowledge of distribution of the non-negative dependent variable, ii) providing more natural way to deal with zero values of the dependent variable, iii) dealing better with sources of heterogeneity within larger panel-type dataset instead to resort to log-linear regressions (Correia et al., 2019), iv) allowing flexibility with multiple fixed effects and interactions (Fally, 2015), v) Correia's STATA command exploits build-in packages which allows for faster estimation of parameters of interest even in presence of multiple fixed effects, by dropping problematic observations to avoid multicollinearity (Santos Silva and Tenreyro, 2010). In fact, recent article of Correia, Guimaraes, and Zylkin (2019) discusses the necessary and sufficient conditions for the existence of estimates in a wide class of GLM models and show that, in the case of Poisson regression, it is always possible to find MLE estimates if some observations are dropped from the sample. Hence, they promote the Pseudo-Poisson Maximum Likelihood as valid tool that can easily detect and discard separated observations that do not convey relevant information for the estimation process. Finally, the Pseudo Poisson takes care to check for existence of maximum likelihood results and introduces promising concepts for accelerating nonlinear estimation with high-dimensional covariates (Correia et al., 2019).

The following section reports results obtained from the estimation of Equation (2) using Zero-Inflated Poisson and Pseudo Poisson Maximum Likelihood with multiple High Dimensional Fixed Effects using the STATA build-in command of ZIP (Greene, 2012; Long, 1997) and not-build in command of PPMLHDFE (Correia et al., 2020), respectively.

IV. Results

We perform the Zero-Inflated Poisson (ZIP) and Pseudo Poisson Maximum Likelihood with multiple High Dimensional Fixed Effects (PPMLHDFE), and we compare results obtained.

Table 2 reports results obtained through the already mentioned models with gravity specification. In fact, we insert time, in place of usual distance variable, as more accurate measure of distance within provinces, and average population as mass variables of source and destination provinces, respectively. Besides, Table x reports in Notes law deterrence, used as control variable for stability of judicial enforcement, for origin and destination places as well as the fixed effects of Italian Centre and South macro-areas interacting with year of 2010-2017. The decision to insert these fixed effects grouped for macro-area is due to detect common and relevant fixed effect within the same area that cannot be revealed if, instead, provinces are used. Robust cluster per paired provinces are also reported. At the bottom of Table x, statistic report 20.808 total observations, Wald Chi-Square tests, p-values for the Chi-Square and Pseudo R-Squared as measure of goodness of fit for Pseudo-Poisson model. For ZIP models, we use cluster instead of Vuong option, that results in a fairly large change in the model Chi-Square, becoming Wald Chi-Square. This statistic is based on log pseudo-likelihoods instead of log-likelihoods and values obtained on columns I-IV suggest that model specifications are significant. Same results are obtained for PPMLHDFE from the Wald Chi-Square test and its p-values on columns V-VIII. In addition, PPMLHDFE presents pseudo-R-Squared as measure of goodness of fit, being close to the value of perfect fit.

Column (I) reports as count dependent variable the number of enrolled students to university, column (II) uses the number of students enrolled to university courses belonging to the field of Social Science (ERC-1), column (III) reports the number of students enrolled to university courses belonging to the field of Physical Science (ERC-2) and column (IV) uses the number of students enrolled to university courses belonging to the field of Life Science (ERC-3). Even columns (V) to (VIII) display estimates performed via PPMLHDFE with the same already-mentioned count dependent variables of enrolled students. Both models display similarities of results for signs especially. However, PPMLHDFE has better performance than ZIP when the statistical significances of its coefficients are reported.

Variables of time and of average population of origin and destination provinces are statistically significant and have the expected signs: in fact, time exhibits negative sign, meanwhile the average population of source and destination provinces have positive signs. Signs and magnitude of these variables are consistent with the predictions of gravity model: the greater the masses, meaning more services and entertainment activities available for citizens, the greater the number of students attracted from smaller to bigger provinces, while the highest time, expressed as time devoted to travel by car, the lowest the emigration occurrence due to monetary and non-monetary costs of leaving origin places (Dotti et al., 2013).

Then, the main independent variable of total corruption exhibits the expected positive sign for origin provinces and negative sign for destination provinces. Besides, it is strongly significant for ZIP and PPMLHDFE, meaning that an increase of total corruption at origin province positively influences, on average, young brain drains from source province toward destination provinces where corruption is significantly lower. This result is in line with the already-cited study of Poprawe (2015), conducted for 230 OECDs within period 2000-2010, demonstrating that young students are interested to study in areas where merit is rewarded fully. In addition, this finding is confirmed by recent literature on the Italian brain drain case that assesses how bad environments are found to have detrimental effects over young brain retaining within Italian regions (Nifo and Vecchione, 2014; Ciriaci, 2013; Dotti et al., 2013).

However, this research put forward another issue that has not been widely treated so far, from macroeconomic perspective, by recent research literature for the Italian case and our study is a first attempt to highlight the variation of sensibility to corruption among young skilled people and more in general understanding whether the factors at play in the decision to migrate are homogeneous across students of different ERC macro-area of studies. Thus, sensibility to corruption varies among students who decided to apply to different study-courses at university: in fact, our results suggest that students who decide to enrol for courses belonging to fields of Social Sciences (ERC-1) and Physical Sciences (ERC-2) are more sensitive to corruption at origin province (which is statistical significant at 5% and 10% confidence level on columns II and III, significant at 5% and 2% on columns VI and VII) than students who enrol to courses belonging to the field of Life Science (ERC-3) and for whom corruption at origin province is not statistically significant (columns IV and VIII) for deciding to rest or to move to universities of different destination provinces. Hence, sensitivity of corruption varies according to the study interests of students: young students who apply for Law, Economics, Political Science or Engineering courses have higher self-awareness on society-issues and are more sensitive to corruption than those students who apply for Medicine and are interested on health-issues only.

Then, the coefficient of real GDP per capita has the expected positive sign for origin province and negative for destination province. For ZIP model, the real GDP per capita is not statistically significant neither for origin nor for destination provinces. For PPMLHDFE it is statistically meaningful at origin province (at 5% and 10% confidence level on columns V, VI and VII) and at destination province seldomly (10% confidence on column VII). Hence, real GDP per capita allows to control for wealth effects: higher well-being status permits to bear higher educational investment costs and incentivize skilled people to leave their origin place because costs of studying abroad become affordable as well as to rest in their origin place if it is rich and it provides good quality of public services, as higher education due to the presence of better universities.

Besides, employment rate reveals the expected negative sign at origin province and the expected positive sign for destination province. It is statistically meaningful for origin province via ZIP and PPMLHDFE, while is statistically significant for destination province via PPMLHDFE only (on columns V VI and VII). Intuitively, an increase in employability of young skilled individuals in source province, on average, refrain them to tackle decision to move away from home and, if they move away, they will choose a destination province that gives them good chances of employability after their study period at university. Thus, students' migration responds to labour market incentives also. This result seems to confirm the hypothesis that students make their migration choices by observing current labour outcomes in terms of job opportunities offered by the place of interest (Dotti et al., 2013).

Then, size of university is inserted and has the expected negative sign for origin provinces and the expected positive sign for destination provinces. Size of university is statistically significant for origin and destination provinces via PPMLHDFE (1% and 10% confidence level on columns V, VI, VII and VIII). Thus, if the size of university of source province is large, the number of resident students who decide to move to different universities of destination provinces decreases. In fact, students prefer to attend large universities because offer greater number of didactic and international courses available than those offered by smaller universities. The increase of enrolled students permits to large universities to become even larger because the development of additional courses is incentivized.

Moreover, we added another variable related education that is the standardized quality of university. As stated in section III, this variable is derived by own elaboration and considers the current values that better assess the quality of universities from Italian students' perspective, contrary to the commonly used indicators, offered by CENSIS and/or Sole 24 Ore sources. In fact, CENSIS indicators consider parameters out of the real interests of students once they get graduation, such as infrastructures, scholarships, web presence, quality of research and internationalization. As expected, the sign of quality of university is negative for origin province and positive, as PPMLHDFE model suggest, for destination provinces. Quality of university is statistically significant for origin provinces (1% confidence level on columns V, VI and VIII) and seldom significant for destination provinces (5% and 10% on columns VII and VIII). This suggests that higher quality of university at origin province, in terms of achieving degree on time, having good opportunities to find a job immediately after graduation and earning reasonable income, tends to reduce young skilled emigration to different universities. This result is also confirmed by Dotti et al. (2013), who demonstrate that young students prefer to attend universities located in provinces with dynamic labour market because represents a source of opportunities for faster development of their professional career.

Then, standardized quality of life is added and elaborated with the same method with which we have formulated the standardized quality of university, already presented in section III. The sign of quality of life is negative, as expected, for origin provinces and positive, as expected, for destination provinces with small exceptions. Although it is seldom significant (on columns III and VII) for origin and destination provinces, an increase of living standards is associated, on average, with fewer skilled individuals who migrate from origin to different provinces. In fact, skilled individuals are looking forward to living in communities that offer to them efficient public services (childcare and elderly care as well as healthcare) and safeguard their safety and security. This result is also in line with studies conducted by Beine et al, 2014; Nifo and Vecchione, 2014; Ciriaci, 2013, Dotti et al., 2013; Van Bouwel and Veugelers 2013, supporting the thesis that overall better standards of living retain or/and attract skilled people.

Furthermore, dummies for infrastructures indicating the presence of airport, port and high-speed railway station at origin and destination provinces are added to control whether their presence ease the students' mobility within provinces. Although we have already mentioned time as measurement of distance viable by car, adding these dummies permit to have more information related viability within provinces. Airport dummy displays the expected positive sign for origin province and the expected positive sign for destination province. It is seldom significant bur its presence ensures that high-distanced provinces can be reached in less hours than hours travelled by car otherwise. Similar evidence is obtained in the case of high-speed railway and port dummy. Ports are essential logistic infrastructures needed for connecting minor and major islands to mainland, preventing them from isolation. In addition, port facilitates transport by car and by train and does not discourage mobility of individuals.

Overall, these two models maintain same signs of variables used but the statistical significance of coefficients appears more frequent when PPMLHDFE is used instead of ZIP. Hence, using PPMLHDFE has permitted us to gain decisive information than that offered by ZIP. Although PPMLHDFE is more complex, its use resulted to be worthwhile because it returns robust results and easily manage multiple high dimensional fixed effects. However, for sake of completeness, it seemed appropriate to compare its results with the ones provided by ZIP.

__Table 2__. Estimated main results with ZIP and PPMLHDFE.

	Zero-Inflated Poisson	**Zero-Inflated Poisson**	**Zero-Inflated Poisson**	**Zero-Inflated Poisson**	**Pseudo-Poisson Max. Likelihood**	**Pseudo-Poisson Max. Likelihood**	**Pseudo-Poisson Max. Likelihood**	**Pseudo-Poisson Max. Likelihood**
	I	II	III	IV	V	VI	VII	VIII
	Enrolled	**ERC-1**	**ERC-2**	**ERC-3**	**Enrolled**	**ERC-1**	**ERC-2**	**ERC-3**
time	-.02225***	-.02284***	-.01729***	-.0151***	-.02673***	-.02846***	-.02593***	-.02423***
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
population	2.58e-07**	3.46e-07***	3.46e-07***	7.54e-08	3.03e-07***	3.77e-07***	2.65e-07**	1.79e-07
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
population_j	3.64e-07***	2.52e-07***	2.95e-07**	5.01e-07***	2.83e-07***	1.88e-07**	3.47e-07***	4.10e-07***
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
corruption (tot)	.001469**	.0009438*	.001603**	.0005472	.001951***	.001288**	.003353***	.00101
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
corruption_j (tot)	-.00311***	-.002182***	-.003786***	-.002017***	-.003373***	-.00231***	-.0055***	-.002443***
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
real GDP per capita	.0000138	.0000127	-7.66e-06	4.90e-06	.0000214**	.0000196*	.0000225*	.0000205
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
real GDP per capita_j	-8.19e-06	3.11e-06	1.10e-06	-.0000114	-.0000104	3.82e-06	-.0000232*	-.0000254
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
employment	-.0469***	-.03719***	-.03909***	-.03518**	-.05837***	-.05432***	-.05377***	-.07205***
	{0.013}	{0.013}	{0.014}	{0.014}	{0.013}	{0.013}	{0.015}	{0.014}
employment_j	.02083	.006165	.009782	.01051	.03291**	.02658*	.02968	.0511***
	{0.014}	{0.015}	{0.016}	{0.015}	{0.014}	{0.015}	{0.018}	{0.016}
size university	-.1436*	-.1179	-.05357	.002883	-.2933***	-.2512***	-.3805***	-.2128*
	{0.086}	{0.084}	{0.105}	{0.108}	{0.081}	{0.081}	{0.099}	{0.115}
size university_j	.3751***	.2002**	.5945***	.2171*	.5308***	.3506***	.8872***	.458***
	{0.095}	{0.091}	{0.124}	{0.121}	{0.091}	{0.088}	{0.112}	{0.137}
standardized quality university	-.3936**	-.3704**	-.263	-.5743***	-.4515***	-.5002***	-.256	-.6431***
	{0.164}	{0.164}	{0.214}	{0.192}	{0.157}	{0.168}	{0.201}	{0.234}
standardized quality university_j	-.04433	-.01318	-.4454*	.267	.09124	.2046	-.4224*	.5262**
	{0.173}	{0.172}	{0.256}	{0.217}	{0.154}	{0.168}	{0.216}	{0.219}
standardized quality life	-.421	-.329	-.6991**	-.2478	-.3295	-.1113	-.7565**	-.2235
	{0.279}	{0.267}	{0.305}	{0.357}	{0.277}	{0.262}	{0.310}	{0.416}
standardized quality life_j	.2607	-.01667	.5656**	.1987	.2077	-.241	.8157***	.4393
	{0.240}	{0.262}	{0.272}	{0.311}	{0.235}	{0.258}	{0.265}	{0.341}
Dairport	.045	.03806	.04933	.1911	.04778	.05861	.04842	.03189
	{0.116}	{0.107}	{0.136}	{0.153}	{0.113}	{0.114}	{0.135}	{0.158}
Dairport_j	.1044	.1167	-.008307	.3003*	.08623	.05885	-.05277	.3503*
	{0.131}	{0.119}	{0.166}	{0.154}	{0.130}	{0.126}	{0.157}	{0.182}
DTAV	-.0616	-.1039	.04328	.1975	-.2519	-.2952*	-.3534	-.03559
	{0.192}	{0.185}	{0.232}	{0.200}	{0.172}	{0.162}	{0.220}	{0.198}
DTAV_j	.1386	.2066	-.01834	-.2797	.4581***	.5152***	.5659***	.1758
	{0.180}	{0.174}	{0.232}	{0.232}	{0.151}	{0.146}	{0.187}	{0.195}
Dport	.283**	.2893**	.2047	.1712	.3582***	.403***	.2298	.4739***
	{0.124}	{0.140}	{0.154}	{0.160}	{0.121}	{0.138}	{0.143}	{0.161}
Dport_j	-.04648	-.09002	.1137	-.118	-.08843	-.173	.2069	-.325
	{0.147}	{0.161}	{0.194}	{0.193}	{0.151}	{0.166}	{0.190}	{0.206}
N.	20808	20808	20808	20808	20808	20808	20808	20808
Wald Chi Square test	545.44	597.845	367.115	202.422	364.188	375.918	242.792	236.564
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pseudo-R2					0.8725	0.8883	0.8305	0.8067

Notes: For both model's specifications, we inserted law deterrence variables for origin and destination provinces as control values. In addition, fixed effects of dummies for macro-areas of North, Centre and South of origin and destination are included with interaction with year (2010-2017 period). Standard errors account for clustering of provinces.

***, ** and * denote coefficients that are statistically significant at 1%, 5% and 10%, respectively.

Robustness Check

The results in the previous section seem to confirm our hypothesis that higher level of corruption at origin province positively influences students' migration and that sensibility to corruption varies among students and it is significant for students who are enrolled to academic courses that belong to Social and Physical Sciences but not for those who attend courses belong to Life Science. In this section, we discuss two robustness checks that we have performed.

Our methodology proves to be robust if we consider i) the addition of the interaction term of distance and transport infrastructures in the regressions, ii) the introduction of the macro area specification that accounts for long-distance skilled movements from South to Centre-North of Italy iii) the addition of the restriction related the movement of students from origin to different destinations without accounting for those who remain in origin province.

Table 3 presents the interaction terms of the dummies of transport infrastructures with distance, expressed in kilometres (km). Interestingly, all values are statistically significant and indicates that their presence reduce the distance (km) of travelling between provinces (as its negative sign mainly suggests). In fact, provinces with airports are more attractive because renders distances between provinces, located far away, achievable in less time. The same condition is provided by the interaction terms of distance with high-speed train stations, confirming that their presence incentivizes mobility because permit faster connections by reducing distances within different provinces. On the other hand, the interaction term of distance and ports presents positive sign for origin and negative sign for destination provinces. As evidence suggests, connections by ports are not faster as airports and trains to reduce the distances between provinces. Hence, cities with ports exercise low attractiveness to fast mobility. Besides, Table 3 shows-off that the sign of our main independent variable of total corruption is preserved although loses, in certain cases, statistical significance. At the same time, the sensibility to corruption of students enrolled to courses belonging to different study-fields is in line, as demonstrated via PPMLHDFE on columns V-VIII of Table 4, with our previously discovered results.

For the remaining variables of Table 3, even though part of these of become somewhat more unstable (for example, real GDP per capita loses significance in all cases for origin and destination provinces, size of university in three cases for origin provinces and standardized quality of university in sixth cases for origin and destination provinces), the sign of all the estimated coefficients is overall preserved.

Furthermore, Table 4 contains ZIP and PPMLHDFE regressions with the specification of movements between macro-area. We check if corruption has positive influence when we consider migration between provinces belonging to non-adjacent macro-area. To evaluate the long-distance skilled movements, we group the Italian provinces into three macro-areas, namely the North, the Centre and the South. This categorical variable, designed for origin and destination j, assumes the values of 1 for identifying the North, of 2 for the Centre and of 3 for the South. For all regressions (columns I to VIII), we insert the condition of skilled movements by using as benchmark the value of the Centre (if macro-area ≥2 & macro-area_j ≦2) for indicating the skilled flow from the South toward Centre-North. Then, we get almost same results showed by the regressions mentioned in the above robustness checks. For ZIP and PPMLHDFE, time and population variables preserve their expected signs, and they are statistically meaningful. The main independent variable of total corruption maintains both positive sign and statistical significance for origin provinces as well as negative sign and statistical significance for destination provinces in both specifications. At the same time, the sensibility to corruption of students enrolled to courses belonging to different study-fields is in line with signs and statical significance as expected in both statistical models. Hence, this exercise confirms that corruption exercises positive influence over long-distance skilled movements.

For the remaining variables of Table 4, even though part of these become somewhat more unstable with ZIP (for example, real GDP per capita loses significance in all cases for origin and destination provinces, size of university in fourth cases for origin provinces, standardized quality of university in third cases for origin and third cases for destination provinces, standardized quality of life in fourth cases for origin and third cases for destination provinces and large part of dummy related airport and port maintain signs but lose statistical significances), rather than the results presented by PPMLHDFE that are more stable because they not only preserve the signs of the estimated coefficients but also exhibit statistical significance at 10% and 5% for majority part of the already described variables of interest.

Finally, table 5 uses the restriction of evaluating only those people who migrate and do not consider origin as destination provinces. Population, distance, economic and quality of life and university variables maintain their expected signs and their statistical significance. However, our main independent variables become somewhat unstable. Although sign and statistical significance of corruption at destination places are preserved, sign and significance for corruption at origin province are lost.

To sum up, even with different specifications, signs and the statistical significance of the variables used are preserved and confirm the robustness of results obtained from our analysis (Table 2). In addition, the robustness check highlights the statistical power of PPMLHDFE, among the Poisson Family, as an ideal model to be used in gravity set-up.

Table 3. *Robustness Check with interaction term of distance and infrastructure*

	Zero-Inflated Poisson	**Zero-Inflated Poisson**	**Zero-Inflated Poisson**	**Zero-Inflated Poisson**	**Pseudo-Poisson Max. Likelihood**	**Pseudo-Poisson Max. Likelihood**	**Pseudo-Poisson Max. Likelihood**	**Pseudo-Poisson Max. Likelihood**
	I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*
	Enrolled	**ERC-1**	**ERC-2**	**ERC-3**	**Enrolled**	**ERC-1**	**ERC-2**	**ERC-3**
time	-.04807***	-.04631***	-.04019***	-.04538***	-.04739***	-.04522***	-.04314***	-.0523***
	{0.006}	{0.006}	{0.010}	{0.006}	{0.006}	{0.006}	{0.011}	{0.007}
population	2.96e-07***	4.19e-07***	3.18e-07***	1.52e-07	3.09e-07***	4.55e-07***	1.58e-07	1.78e-07
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
population_j	2.30e-07***	1.01e-07	2.00e-07**	3.74e-07***	2.36e-07***	8.35e-08	3.76e-07***	3.79e-07***
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
corruption	.0004787	.0000645	.0006148	-.0001855	.001133*	.0004493	.002511***	.0005969
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
corruption_j	-.001822***	-.001111**	-.002444***	-.001049	-.002345***	-.001291**	-.004413***	-.001719**
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
real GDP per capita	-.0000117	-.0000119	-.0000219*	-.0000191	-.0000117	-.0000182	-3.94e-06	-.0000122
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
real GDP per capita_j	.0000148	.0000255**	.0000161	5.90e-06	.0000175	.0000367***	-1.04e-06	1.69e-07
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
employment	-.02789**	-.0207	-.0169	-.01051	-.03515***	-.02854**	-.03443**	-.04897***
	{0.013}	{0.013}	{0.014}	{0.015}	{0.013}	{0.013}	{0.014}	{0.017}
employment_j	.006193	-.00674	-.006875	-.005997	.0136	.003827	.01241	.03414*
	{0.013}	{0.013}	{0.016}	{0.016}	{0.013}	{0.014}	{0.017}	{0.018}
size university	-.1986**	-.1655*	-.2351***	-.03884	-.3281***	-.2756***	-.4881***	-.1888
	{0.090}	{0.091}	{0.090}	{0.111}	{0.085}	{0.088}	{0.090}	{0.133}
size university_j	.3901***	.2182**	.6791***	.1894	.5242***	.3428***	.9566***	.3812**
	{0.100}	{0.096}	{0.112}	{0.125}	{0.099}	{0.097}	{0.116}	{0.158}
standardized quality university	-.5734**	-.5638**	-.3478	-.6193**	-.6981***	-.7616***	-.4729*	-.8689***
	{0.230}	{0.222}	{0.282}	{0.242}	{0.211}	{0.210}	{0.263}	{0.291}
standardized quality university_j	.06558	.1266	-.343	.2868	.2472	.386*	-.269	.6246**
	{0.225}	{0.216}	{0.301}	{0.258}	{0.194}	{0.201}	{0.256}	{0.260}
standardized quality life	-.4633*	-.3041	-.7904***	-.2572	-.4385	-.1699	-.9315***	-.3655
	{0.279}	{0.282}	{0.296}	{0.342}	{0.276}	{0.278}	{0.310}	{0.410}
standardized quality life_j	.3069	-.05663	.6693**	.2851	.3226	-.1733	.9733***	.6201*
	{0.253}	{0.285}	{0.268}	{0.315}	{0.250}	{0.282}	{0.272}	{0.368}
1.Dairport#c.dist	-.002864***	-.002638***	-.002057***	-.002624***	-.003034***	-.002775***	-.002864***	-.003528***
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
1.Dairport_j#c.dist	.002299*	.001325	.003318***	-.0001714	.002382*	.001418	.005683***	.0001322
	{0.001}	{0.002}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.002}
1.DTAV#c.dist	-.007416***	-.007973***	-.00646***	-.004531**	-.008064***	-.008791***	-.008332***	-.006684***
	{0.002}	{0.002}	{0.001}	{0.002}	{0.002}	{0.002}	{0.002}	{0.002}
1.DTAV_j#c.dist	.001957*	.003265**	.0006851	.0008236	.00516***	.007292***	.003464***	.004213***
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
1.Dport#c.dist	.004399***	.004536***	.003453***	.002825***	.005443***	.005883***	.005302***	.004756***
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
1.Dport_j#c.dist	-.005315***	-.004289**	-.008261***	-.002481*	-.004747***	-.003487**	-.00959***	-.002258
	{0.002}	{0.002}	{0.002}	{0.001}	{0.002}	{0.002}	{0.002}	{0.001}
N.obs	20808	20808	20808	20808	20808	20808	20808	20808
Wald chi2 test	1110.97	10866.70	8692.09	6128.40	6224.91	6895.78	3952.55	4292.53
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pseudo-R2					0.9102	0.9207	0.8833	0.8396

Notes: For both model's specifications, we inserted law deterrence variables for origin and destination provinces as control values. In addition, fixed effects of dummies for macro-areas of North, Centre and South of origin and destination are included with interaction with year (2010-2017 period). Standard errors account for clustering of provinces.

***, ** and * denote coefficients that are statistically significant at 1%, 5% and 10%, respectively.

Table 4. *Robustness Check with brain flow from South to Centre-North*

	Zero-Inflated Poisson	**Zero-Inflated Poisson**	**Zero-Inflated Poisson**	**Zero-Inflated Poisson**	**Pseudo-Poisson Max. Likelihood**	**Pseudo-Poisson Max. Likelihood**	**Pseudo-Poisson Max. Likelihood**	**Pseudo-Poisson Max. Likelihood**
	I	II	III	IV	V	VI	VII	VIII
	Enrolled	**ERC-1**	**ERC-2**	**ERC-3**	**Enrolled**	**ERC-1**	**ERC-2**	**ERC-3**
time	-.01109***	-.01196***	-.007201***	-.006969***	-.01519***	-.01787***	-.0134***	-.01401***
	{0.002}	{0.002}	{0.002}	{0.002}	{0.002}	{0.002}	{0.002}	{0.002}
population	4.45e-08	-2.23e-09	1.78e-07	6.15e-08	2.17e-08	-1.80e-08	7.86e-08	1.61e-07
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
population_j	6.59e-07**	5.59e-07**	5.21e-07**	7.53e-07***	8.63e-07***	7.15e-07***	9.84e-07***	8.70e-07***
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
corruption	.002295**	.002142*	.002016*	.001052	.002933***	.002663**	.003606***	.000986
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
corruption_j	-.002598**	-.0004577	-.00346***	-.002794**	-.004281***	-.001389*	-.007296***	-.002392**
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
real GDP per capita	.0000512	.000068	2.40e-06	-.0000169	.0000644	.0000705	.0000877	.0000196
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
real GDP per capita_j	-.0000366	-2.95e-06	-.0000483*	-.0000325	-.000018	.0000468*	-.0001072***	.0000505
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
employment	-.05295**	-.06446**	-.03964	.005224	-.06342**	-.07905***	-.05653*	-.05437*
	{0.026}	{0.026}	{0.027}	{0.022}	{0.028}	{0.028}	{0.031}	{0.029}
employment_j	.06794	.0618	.02415	.05548	.06271	.0729*	.04156	.07148*
	{0.046}	{0.042}	{0.035}	{0.040}	{0.040}	{0.041}	{0.044}	{0.038}
size university	-.1471	-.09926	.1235	.0003323	-.4015**	-.3617*	-.5229**	-.2718
	{0.198}	{0.211}	{0.189}	{0.169}	{0.189}	{0.201}	{0.204}	{0.174}
size university_j	.1024	-.102	.2694	-.1101	.3091	.008161	.7892***	-.004971
	{0.239}	{0.231}	{0.262}	{0.251}	{0.226}	{0.220}	{0.235}	{0.230}
standardized quality university	-.4872	-.2808	-.2667	-1.114***	-.5163	-.5007	-.2516	-1.128**
	{0.404}	{0.445}	{0.368}	{0.346}	{0.416}	{0.450}	{0.433}	{0.454}
standardized quality university_j	-.03791	-.09394	-.9129*	-.6405	.6538*	.9994**	.3474	.4144
	{0.492}	{0.492}	{0.491}	{0.461}	{0.361}	{0.420}	{0.345}	{0.354}
standardized quality life	-.7992	-.4636	-.8644	-.4922	-.8757	-.4774	-1.412**	-.3702
	{0.614}	{0.584}	{0.608}	{0.551}	{0.535}	{0.518}	{0.579}	{0.561}
standardized quality life_j	.7041	.0855	.707	1.506***	.3916	-.1663	.8206**	1.527***
	{0.430}	{0.444}	{0.456}	{0.422}	{0.339}	{0.380}	{0.371}	{0.342}
Dairport	.04641	.01021	.2715	.5028	.05664	.05363	.05942	-.01651
	{0.314}	{0.310}	{0.277}	{0.311}	{0.314}	{0.310}	{0.336}	{0.324}
Dairport_j	.4811	.1882	.6754*	-.05106	.907***	.4923	1.574***	.7017**
	{0.354}	{0.377}	{0.357}	{0.343}	{0.330}	{0.344}	{0.344}	{0.321}
DTAV	-.4774	-.5029	-.3614	.1536	-.6425	-.6914	-.7786	-.4617
	{0.570}	{0.510}	{0.588}	{0.456}	{0.591}	{0.509}	{0.733}	{0.636}
DTAV_j	.4008	.3239	.5708	.06311	.4189	.3958	.6328	-.1617
	{0.576}	{0.543}	{0.629}	{0.541}	{0.470}	{0.467}	{0.496}	{0.582}
Dport	.3218	.5675*	-.005406	-.3724	.7208**	1.014***	.4679	.5285
	{0.331}	{0.337}	{0.327}	{0.322}	{0.342}	{0.350}	{0.418}	{0.385}
Dport_j	-.4123	-.9637**	.08335	.2207	-1.153***	-1.688***	-.7224*	-1.097**
	{0.406}	{0.377}	{0.393}	{0.425}	{0.418}	{0.421}	{0.385}	{0.441}
N	7440	7440	7440	7440	7440	7440	7440	7440
Wald Chi Square Test	314.527	395.558	245.512	210.900	853.81	743.15	781.21	693.74
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pseudo-R2					0.7947	0.8252	0.7397	0.7408

Notes: For both model's specifications, we inserted law deterrence variables for origin and destination provinces as control values. In addition, fixed effects of dummies for macro-areas of North, Centre and South of origin and destination are included with interaction with year (2010-2017 period). Standard errors account for clustering of provinces.

***, ** and * denote coefficients that are statistically significant at 1%, 5% and 10%, respectively.

Table 5. *Robustness Check that does not account origin as destination province*

	Zero-Inflated Poisson	Zero-Inflated Poisson	Zero-Inflated Poisson	Zero-Inflated Poisson	Pseudo-Poisson Max. Likelihood	Pseudo-Poisson Max. Likelihood	Pseudo-Poisson Max. Likelihood	Pseudo-Poisson Max. Likelihood
	I **Enrolled**	*II* **ERC-1**	*III* **ERC-2**	*IV* **ERC-3**	*V* **Enrolled**	*VI* **ERC-1**	*VII* **ERC-2**	*VIII* **ERC-3**
time	-.009159***	-.009144***	-.005104***	-.006103***	-.01486***	-.01651***	-.01309***	-.01447***
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.002}	{0.002}	{0.002}
population	2.83e-07*	2.92e-07*	3.22e-07*	1.72e-07	4.08e-07**	4.23e-07**	3.67e-07	4.54e-07**
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
population_j	2.89e-07***	1.14e-07	2.17e-07**	4.38e-07***	2.73e-07***	8.66e-08	3.90e-07***	4.96e-07***
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
corruption	-.001256	-.001715	-.001568	-.001196	-.0007461	-.001129	-.0005567	-.0003865
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.002}	{0.001}
corruption_j	-.002326***	-.00156*	-.0034***	-.000917	-.002472**	-.001121	-.00575***	-.0006962
	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}	{0.001}
real GDP per capita	.0000314	.0000333*	2.41e-06	.0000204	.0000649***	.0000692***	.000054*	.0000681***
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
real GDP per capita_j	.0000276	.0000356*	.0000224	.0000274	.0000412**	.0000675***	.0000202	.000025
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
employment	-.04689**	-.03719*	-.02093	-.03271*	-.07583***	-.06739***	-.07136***	-.09384***
	{0.021}	{0.022}	{0.025}	{0.018}	{0.020}	{0.021}	{0.026}	{0.020}
employment_j	.002639	-.006994	-.01987	-.01417	.01423	.01572	-.007304	.02807
	{0.019}	{0.019}	{0.024}	{0.020}	{0.021}	{0.020}	{0.028}	{0.024}
size university	-.1584	-.1005	-.1676	-.1026	-.3109***	-.2265*	-.4592***	-.2858**
	{0.115}	{0.122}	{0.121}	{0.100}	{0.116}	{0.123}	{0.136}	{0.132}
size university_j	.2764**	.07778	.6028***	.07427	.5026***	.31**	.9318***	.3946*
	{0.131}	{0.129}	{0.142}	{0.152}	{0.148}	{0.153}	{0.170}	{0.216}
standardized quality university	-.6174**	-.682**	-.3053	-.7393**	-.6858**	-.8083***	-.3035	-.9734**
	{0.293}	{0.272}	{0.346}	{0.313}	{0.278}	{0.255}	{0.326}	{0.400}
standardized quality university_j	-.1615	-.2456	-.3854	.1414	.07268	.1022	-.259	.4464
	{0.283}	{0.252}	{0.392}	{0.336}	{0.254}	{0.282}	{0.332}	{0.290}
standardized quality life	-.4232	-.3952	-1.052**	-.07948	-.1689	.06616	-.8779**	.3297
	{0.395}	{0.426}	{0.428}	{0.326}	{0.344}	{0.366}	{0.396}	{0.450}
Standardized quality life_j	.4152	.2126	.7623**	.1232	.3759	-.1658	1.183***	.5042
	{0.316}	{0.376}	{0.329}	{0.339}	{0.299}	{0.364}	{0.356}	{0.430}
Dairport	-.1755	-.1565	-.09282	-.1511	-.231	-.2276	-.1414	-.3809*
	{0.198}	{0.189}	{0.240}	{0.172}	{0.194}	{0.197}	{0.249}	{0.208}
Dairport_j	-.2772	-.2403	-.4601**	-.1323	-.3113	-.3565*	-.4231*	-.05077
	{0.182}	{0.163}	{0.207}	{0.175}	{0.204}	{0.197}	{0.257}	{0.245}
DTAV	-.7219*	-.7111*	-.5057	-.4255	-1.083**	-1.025**	-1.103**	-1.21**
	{0.403}	{0.385}	{0.395}	{0.318}	{0.435}	{0.410}	{0.516}	{0.486}
DTAV_j	.05579	.1794	.09196	-.5101**	.4732*	.5596**	.671**	-.1133
	{0.256}	{0.234}	{0.269}	{0.259}	{0.246}	{0.221}	{0.314}	{0.292}
Dport	.3308	.3737	.0765	.3232	.524*	.6464**	.2977	.6472**
	{0.262}	{0.255}	{0.284}	{0.233}	{0.279}	{0.275}	{0.333}	{0.289}
Dport_j	.1113	.1432	.3765	-.0001862	.09573	.19	.2683	-.2951
	{0.208}	{0.210}	{0.247}	{0.218}	{0.260}	{0.272}	{0.355}	{0.296}
N	20400	20400	20400	20400	20400	20400	20400	20400
Wald Chi Square test	298.16	217.26	180.75	355.88	737.40	637.87	627.46	706.23
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pseudo-R2					0.6036	0.6237	0.5193	0.5493

Notes: For both model's specifications, we inserted law deterrence variables for origin and destination provinces as control values. In addition, fixed effects of dummies for macro-areas of North, Centre and South of origin and destination are included with interaction with year (2010-2017 period). Standard errors account for clustering of provinces.

***, ** and * denote coefficients that are statistically significant at 1%, 5% and 10%, respectively.

Endogeneity

A critical concern in empirical analysis is endogeneity, which causes biased and inconsistent estimates. While literature on dynamic panel data models offers solutions to deal with endogeneity frequently, for gravity models this issue is seldom threated. Hence, this paper faces the concern of inverse causality that may occur between corruption and inter-provincial skilled migration and offers its scientific contribution on the procedure to deal with endogeneity within gravity set-up.

In the context of gravity models, the common approach used to mitigate the effect of endogeneity relies on the introduction of lagged variables in the regressions. Dotti et al., 2013, who studied how quality of universities and the local labour market conditions in the destination places affected students' mobility behaviour, used lagged explanatory variables to mitigate the risks of endogeneity. The lagged explanatory variables were those which students should had already observed before taking a decision on enrolment. Besides, Arpaia et al., 2018, in examining the bilateral trade in the European area affected by macro-economic variables such as population, real GDP per capita and unemployment, they introduced the lagged exogeneous variables of population at origin, GDP per capita and unemployment for considering reverse causality. Furthermore, in the context of bilateral trade, Ghodsi, 2019, introduced lagged explanatory variables, one-year lag to reduce the endogeneity bias of trade. Besides, Anderson and Yotov, 2020, used two specifications: first, created an instrument for lagged trade by using a restricted form of gravity model that only included the standard gravity variables, which were exogenous. Then, they used the second to fifth lags of the newly constructed trade variable as instruments for the lagged dependent variable in the unrestricted gravity specification. Gu et Shen, 2020, constructed an eigenvector spatial filtering (ESF) hurdle gravity model (ESF-HGM) to examine the determinants of China's skilled and less-skilled internal migrations between 2010 and 2015.

Moreover, another strategy frequently used to deal with endogeneity was the Instrumental Variable (IV) regression. In fact, Biagi et al., 2010, analysed Italian labour mobility into short-distance and long-distance migration due to economic variables, social capital and quality of life variables. In doing so, they estimated these different types of flows using a negative binomial model, augmented with three instruments (football team of destination country, industry employment rate and presence of ATM machines per 1,000 inhabitants) to control for potential endogeneity.

In addition, Bergstrand et al, 2015, examined the effects of economic integration agreements, international borders, and bilateral distance in international trade. To avoid biased results, they implemented panel pair fixed effects, that captured the cross-sectional negative impact of bilateral distance on trade flows.

D'Ambrosio et al., 2018, studied whether migrants promote co-inventorship between regions and foreign countries, and if the social capital of their respective communities favours such innovation networking. They addressed possible endogeneity by issuing more stringent fixed effects to their Pseudo-Poisson Maximum Likelihood Model (PPML).

Zhang, 2020, analysed the role of migrants' taste in international trade and adopted two sets of estimations based on level and difference equations with two instruments. Specifically, he started with estimating the gravity equation in level by two-stage least square method. Then, he further estimated the gravity equation by taking first differences between the two cross-sections to eliminate any time-invariant omitted variables.

Thus, the method that we propose to control for the potential reverse causality between corruption and skilled emigration present similarities, although it is different in the procedure, with the one proposed by Drivas et al., 2020.

We apply a two-stage residual estimation, the equivalent of two-stage least square (2SLS) for count data (Wooldridge, 2018). In doing so, we regress our endogenous variable of corruption on the instrument, that is one-year lagged value of total corruption, in the first stage, conditional upon the other exogenous variables of the original models and recover the predicted residuals (with dual specification *i* and *j*) of this estimation to plug them into our original model (without excluding our endogenous variable) in the second stage; the inference is based on bootstrapping over all the two-step procedure with 50 replications. Final outcomes are shown in Table 6 for bootstrapping with ZIP in second stage and in Table 7 for bootstrapping with PPMLHDFE in second stage. Both tables present similar results: the sign of the coefficient for total corruption at origin is positive while for total corruption at destination is negative. Same expected signs and significances ere preserved for the residuals. These outcomes prove consistency of already demonstrated results above. The bootstrap standard errors are low, although the one presented by Table 6 are relatively lower than the ones showed by Table 7.

Table 6. Bootstrap Results with 2 stages procedure with ZIP in 2nd stage

	Obs. Coef.	**Bootstrap Std. Err.**	**z**	**P>|z|**	**Normal Based [95% Conf. Interval]**	
r(b_corruption)	0.0015866	0.0019895	0.80	0.425	-0.0023128	0.0054861
r(b_corruption_j)	-0.0047955	0.0013653	-3.51	0.000	-0.0074714	-0.0021197
r(b_corr_residuals)	-0.00004	0.0014907	-0.03	0.979	-0.0029618	0.0028818
r(b_corr_residuals_j)	0.002326	0.0011028	2.11	0.035	0.0001645	0.0044876

N. observations: 20.808; N. of Replications: 50 based on 2.601 cluster in panelid

Table 7. Bootstrap Results with 2 stages procedure with PPML in 2nd stage

	Obs. Coef.	**Bootstrap Std. Err.**	**z**	**P>|z|**	**Normal Based [95% Conf. Interval]**	
r(b_corruption)	0.0016962	0.0019472	0.87	0.384	-0.0021203	0.0055127
r(b_corruption_j)	-0.0053151	0.0013857	-3.84	0.000	-0.008031	-0.0025991
r(b_corr_residuals)	0.0005872	0.0015282	0.38	0.701	-0.002408	0.0035824
r(b_corr_residuals_j)	0.0024111	0.0012473	1.93	0.053	-0.0000335	0.0048558

N. observations: 20.808; N. of Replications: 50 based on 2.601 cluster in panelid

V. **Conclusion**

This paper wants to study why young brain drain occurs within Italian provinces nowadays. We have detected one possible factor of influence that is corruption perceived by young resident students in their source context. Skilled individuals prefer to live in places which exhibit lower levels of corruption and better standards of living than the uncertain better-off conditions offered, through unfairness and illegal means, by corrupted environments.

Hence, we have examined the relationship between corruption and young brain drain via Zero-Inflated Poisson and Pseudo-Poisson Maximum Likelihood models with gravity set-up. We found that higher corruption at origin province is associated, on average, with an increase of young brain drain toward destination provinces, where corruption cases are lower. Furthermore, we have found that sensibility to corruption varies among young students who enrol to courses belonging to different fields of study (Social Science, Physical Science and Life Science): from macroeconomic perspective, students who decide to enrol to Law, Economics and/or Engineering courses are, on average more sensitive to corruption and will tend to emigrate more than students enrolled to Medicine courses. Hence, corruption and/or perception of corruption is detrimental to human capital accumulation and sounds to be main determinant of Italian skilled migration to places where social mobility and higher status quo are equitably achieving.

Furthermore, robustness checks have evidenced that, controlling for interactions terms or for long distance movements and endogeneity, results maintain their stability and statistical power.

Thus, the policy implications are straightforward. Public policies can be crucial in offering control of corruption that would prevent young brain drain from origin provinces. First, policymakers should be mindful of differences of sensibility to corruption of young students and should reinforce law procedures that enhance transparency and reduce information asymmetry by discouraging unethical behaviour with introduction of severe social penalties. Second, this study is performed with aggregated data at provincial level that provide more flexibility in capturing specific differences that regional studies do not grab usually. Besides, studies conducted with this type of data and methods return an accurate analysis on corruption and brain drain occurrences that help policymakers to find appropriate solutions to control for these social issues. In addition, controlling for quality of university by channelling more resources into didactic courses that recognize and reward merit, could be valid policy tool to rein young brain drain from origin province, that has its own local university, and sufficient to limit its negative economic effects. Young brain migration as well as corruption represents urgent government-oriented issues for policymakers to be re-addressed. To this end, such policy corrections could reduce corruption and increase trust that, in turn, would decrease skilled diaspora and its consequent socio-economic inequalities. Such conditions seed grounds for better labour market outcomes by speeding-up the economic process of growth convergence within Italian provinces.

Acknowledgments

We are grateful thank toward dott. Cinzia Pellicanò, consultant at Department Statistics of ISTAT, for being available in providing requested data timely and efficiently.

Notes

Further tables of figures and robustness checks and endogeneity test with more replications are available upon request.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

No special funding was provided in doing this research.

ORCID

Emanuele Millemaci: https://orcid.org/0000-0002-9095-7513

Alessandra Patti: https://orcid.org/0000-0003-1993-3620

References

- Aassve, A., Daniele, G., Le Moglie, M., 2018 “Never Forget the First Time: The Persistent Effects of Corruption and the Rise of Populism in Italy”, http://dx.doi.org/10.2139/ssrn.3280498
- Allen, T., Arkolakis, C., Takahashi, Y., 2020 “Universal Gravity” *Journal of Political Economy*, 128(2):393–433.
- Anderson, J. E., van Wincoop, E., 2003 “Gravity with gravitas: A Solution to the Border Puzzle” *American Economic Review*, 93(1):170–192
- Anderson, J. E., Yotov, Y.V., 2016 “Terms of Trade and Global Efficiency Effects of Free Trade Agreements, 1990–2002” *Journal of International Economics*, 99:279–298
- Anderson J.E., Yotov, Y.V., 2020, “Short run Gravity”, *Journal of International Economics* 126:103-341
- Auer, D., Romer, F., Tjaden, J., 2020 “Corruption and the Desire to Leave Quasi-Experimental Evidence on Corruption as a Driver of Emigration Intentions” *IZA Journal of Development and Migration*, 11:7
- Ariu, A., Docquier, F., Squicciarini M., 2016 “Governance Quality and Net Migration Flow” *Regional Science and Urban Economics*, 60: 238-248
- Arpaia, A., Kiss, A., Palvolgyi, B., Turrini, A., 2018. "The effects of European integration and the business cycle on migration flows: a gravity analysis," *Review of World Economics* Springer vol. 154(4):815-834
- Beine, M., Noel, R., Ragot, L., 2014 “Determinants of the international mobility of students” *Economics of Education Review*, 41:40-54

- Bergstrand, J., Larch, M., Yotov,Y., 2015, "Economic integration agreements, border effects, and distance elasticities in the gravity equation", *European Economic Review,* 78 (C):307-327
- Biagi, B., Faggian, A., McCann, P., 2011 "Long and Short Distance Migration in Italy: The Role of Economic, Social and Environmental Characteristics" *Spatial Economic Analysis,* 6(1):111-131, DOI: 10.1080/17421772.2010.540035
- Bonanno, G., Fiorino, N., Garzarelli, G., Rossi, S., 2020 "Public Guarantee Schemes, Corruption and Gender: a European SME-level analysis", *Applied Economics* DOI: 10.1080/00036846.2020.1798342
- Burger, M., van Oort, F., Linders, G., 2009 "On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation" *Spatial Economic Analysis*, 4(2):167-190, DOI: 10.1080/17421770902834327
- Cameron, A., Trivedi, P., 2010 "Microeconometrics using Stata", *Stata Press MUSR,* 8:18.
- Cameron, A., Trivedi, P., 1998 "Regression Analysis of Count Data" NY: Cambridge Press.
- Cameron, A., 2009 "Advances in Count Data Regression Talk for the Applied Statistics Workshop", http://cameron.econ.ucdavis.edu/racd/count.html
- Ciriaci, D., 2014 "Does University Quality Influence the Interregional Mobility of Students and Graduates? The Case of Italy" *Regional Studies*, 48(10):1592-1608, DOI: 10.1080/00343404.2013.821569
- Cooray, A., Schneider, F., 2016 "Does Corruption promote Emigration? An Empirical Examination" *Journal of Population Economics*, 29:293-310, DOI 10.1007/s00148-015-0563-y

- Corrado, G., Rossetti, F., 2018, “Public Corruption: A Study across Regions in Italy” *Journal of Policy Modelling*, https://doi.org/10.1016/j.jpolmod.2018.01.001
- Correia, S., Guimarães, P., Zylkin, T., 2020 “Fast Poisson estimation with high-dimensional Fixed Effects” *The Stata Journal*, 20(1):95–115
- D’Ambrosio, A., Montresor, S., Parrilli M., Quatraro, M., 2018: “Migration, communities on the move and international innovation networks: an empirical analysis of Spanish regions”, *Regional Studies*, DOI: 10.1080/00343404.2018.1426850
- Docquier, F., Rapoport, H., 2012, “Globalization, Brain Drain and Development” *Journal of Economic Literature*, 50(3):681-730.
- Dotti, N., F., Fratesi, U., Lenzi, C., Percoco, M., 2013 “Local Labour Markets and the Interregional Mobility of Italian University Students” *Spatial Economic Analysis*, 8(4):443-468, DOI: 10.1080/17421772.2013.833342
- Dreher, A., Gassebner, M., 2013 “Greasing the wheels? The impact of regulations and corruption on firm entry” *Public Choice*, 155:413-43
- Drivas, K., Economidou, C., & Karamanis, D., & Sanders, M., 2020, "Mobility of Highly Skilled Individuals and Local Innovation Activity," *Technological Forecasting & Social Change*, 158:120-144
- Fally, T., 2015 “Structural Gravity and Fixed Effects” *Journal of International Economics*, 97, 76–85, https://doi.org/10.1016/j.jinteco.2015.05.005
- Ghodsi, M., 2019, “The impact of Chinese technical barriers to trade on its manufacturing imports when exporters are heterogeneous”, *Empirical Economics*, https://doi.org/10.1007/s00181-019-01690-9
- Gonzales C., Mesanza R., 2011 “The Determinants of International Student Mobility Flows: An Empirical Study on the Erasmus Programme” *Higher Education*, 62:413-430 DOI 10.1007/s10734-010-9396-5

- Greene, W., 2008 "Econometric Analysis" *Boston Pearson*, Boston
- Gu, H., Shen, T., 2021 "Modelling skilled and less-skilled internal migration in China 2010-2015: Application of an eigenvector spatial filtering hurdle gravity approach", *Population, Space and Place,* Vol. 21 Issue 6, https://doi.org/10.1002/psp.2439
- Guimarães, P., Portugal, P., 2010 "A Simple Feasible Procedure to Fit Models with High-Dimensional Fixed Effects" *STATA Journal* 10, 628–649, https://doi.org/10.1177/ 1536867X1101000406.
- Hardin J.W., Hilbe J.M., 2018 "Generalized Linear Models and Extensions" *Fourth Edition STATA Press*, 598 Pages
- Istituto Nazionale di Statistica - Istat.it
- Lisciandra, M., Millemaci, E., 2017 "The Economic Effect of Corruption in Italy: A Regional Panel Analysis" *Regional Studies,* 51(9): 1387-1398, DOI: 10.1080/00343404.2016.1184244
- Long, J., Freese, J., 2013 "Regression Models for Categorical Dependent Variables Using Stata", *Third Edition College Station*, TX: Stata Press.
- Long, J., 1997 "Regression Models for Categorical and Limited Dependent Variable" *Thousand Oaks*, CA: Sage Publications.
- Mayda, A., 2010 "International Migration: a Panel Data Analysis on the Determinants of Bilateral Flows" *Journal of Population Economics*, 23(4):1249-1274
- Ministero dell'Istruzione dell'Università e della Ricerca - Miur
- Nelder, J., Wedderburn, R., 1972 "Generalized Linear Models" *Journal of the Royal Statistical Society Series A (General), 135*(3):370-384, DOI:10.2307/2344614.
- Nifo, A., Vecchione, G., 2014 "Do Institutions Play a Role in Skilled Migration? The Case of Italy" *Regional Studies,* 48(10):1628-1649, DOI: 10.1080/00343404.2013.835799

- Pfaffermayr, M., 2020 “Constrained Poisson Pseudo Maximum Likelihood Estimation of Structural Gravity Models” *International Economics,* 161:188-198
- Pfaffermayr, M., 2020 “Trade Creation and Trade Diversion of Economic Integration Agreements Revisited: A Constrained Panel Pseudo-Maximum Likelihood Approach.” *Review of World Economics,* 156:985–1024
- Poissonnier, A., 2019 “Iterative Solutions for Structural Gravity Models in Panel” *International Economics,* 157:55–67.
- Poprawe, M., 2015 “On the Relationship between Corruption and Migration: Empirical Evidence from Gravity Model of Migration” *Public Choice,* 163:337-354, DOI 10.1007/s11127-015-0255-x
- Santos Silva, J.M, Tenreyro, S., Windmeijer, F., 2015 “Testing Competing Models for Non-Negative Data with Many Zeros” *Journal of Econometric Methods,* 4 (1):29- 46, DOI: 10.1515/jem-2013-0005.
- Santos Silva, J.M, Tenreyro, S., 2011 “Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator” Economics *Letters,* 112:220-222
- Santos Silva, J.M, Tenreyro, S, 2010 “On the Existence of the Maximum Likelihood Estimates in Poisson Regression” *Economics Letters,* 107:310-312
- Santos Silva, J.M, J., Tenreyro, S. 2006 “The Log of Gravity” *Review of Economics and Statistics*, 88(4):641–658
- Van Bouwel, L., Veugelers, R., 2013 “The Determinants of Students Mobility in Europe: The Quality Dimension” *European Journal of Higher Education,* 3(2):172-190
- Williams T.F., 2020 “Review of Probability Distribution for Modeling Count Data”, e-print 2001.04343, arXiv, stat.ME

- Windmeijer, F.A., Santos Silva, J.M., 1997 “Endogeneity in Count Data Models: An Application to Demand for Health Care” *Journal of Applied Economics*, 12:281–294
- Wooldridge, J.M., 2007 “Inverse probability weighted estimation for general missing data problems” *Journal of Econometrics*, 141:1281–1301
- Wooldridge, J.M, 2018 “Control Function Methods in Applied Econometrics” *The Journal of Human Resources*, *50*(2), 420–445
- Zhang, P., 2020 “Home-biased gravity: The role of migrant tastes in international trade”, *World Development* 129:104863