INITIATIVE ON
Fragility, Conflict, and Migration

CGIAR

Internal Migration Propensity Index (i-MPI) for Guatemala and Honduras

A tool to estimate the household-level likelihood of domestic migration

Francisco Ceballos, Manuel Hernandez, Alvaro Espinoza Hernandez

Working paper
December 31st, 2023

Abstract

Worldwide, the large majority of migration occurs internally, with people relocating domestically within their country of origin. From a policy perspective, anticipating such migration decisions can be useful for targeting development programs and interventions, or as a monitoring tool, to capture the migration 'sentiment' among specific groups of interest. This paper describes the construction of an internal Migration Propensity Index (i-MPI) for Guatemala and Honduras, building on the existing MPI to predict external or cross-border migration available for these countries. The i-MPI aims to objectively measure and track the probability of individuals in a household of permanently migrating to a different location within their country, based on a small subset of household indicators and conditions that highly correlate with the (latent) decision to migrate domestically and, combined, best predict internal migration at the household level. The i-MPI avoids questions about migration intentions, which may be culturally sensitive, prone to misreporting, and difficult to monitor over time. Alternatively, it relies on a concise, reliable, and statistically-robust scoring index obtained from indirect indicators highly associated with migration decisions that are simple and easy to collect in the field. The paper describes the data and methodological approach to derive the index and presents the final i-MPI questionnaires, including a set of accompanying tools to implement it in the field.

Keywords: International migration; Scoring index; Migrants; Migration probability; Guatemala

Acknowledgments

This study was funded by the OneCGIAR Initiative on Fragility, Conflict, and Migration (FCM). The funding source had no role in data collection, data analysis and report writing. The contents and conclusions of the study are the responsibility of the authors and do not necessarily reflect the views of the International Food Policy Research Institute (IFPRI), OneCGIAR, or the FCM Initiative.

1. Introduction

Approximately one billion individuals worldwide were migrants in 2020, according to estimates by the United Nations (McAuliffe & Triandafyllidou, 2021). Of these, almost 8 out of 10 were internal migrants, i.e. people relocating domestically within their country of origin, with Latin America and the Caribbean being the region with the highest rate of internal migration, followed by Africa and Asia (Lucas, 2016). By 2040, the combined result of population growth and internal migration is estimated to render 1.6 billion new urban residents worldwide (National Intelligence Council, Strategic Futures Group, 2021).

The literature on internal or domestic migration extends back several decades. Earlier studies regarded it as a purely rational choice, mainly driven by higher real wages in more developed areas (Todaro, 1969; Todaro & Harris, 1970; Williamson, 1988). In this line, Ahluwalia (2000) emphasizes that internal migration is influenced by economic growth, population redistribution, regional development, and economic and social inequality. More recently, de Brauw, Mueller, & Lee (2014) argue that internal migration flows follow structural transformation trends within countries, in which labor typically relocates from agriculture to manufacturing and service sectors due to economic growth. Several empirical studies find that internal migration waves are driven by regional labor prospects, employment opportunities, and access to infrastructure (Yamada, 2010; Ray & Dutta, 2019; Swinkels et al., 2019), while others show that poverty and lack of educational services in rural communities act as key push factors that drive people to migrate (Barrientos & Dejong, 2006; de Brauw & Giles, 2006), although economic migration does not always result in higher living standards (Vargas-Lundius et al., 2018; Joulaei et al., 2023).

More recent evidence links internal migration to climate shocks and conflict. Disruptive climate events, such as natural disasters or extreme temperature fluctuations, can push people to move domestically or across borders into neighboring countries (Ibáñez et al., 2022; Alverio, Sowers, & Weinthal, 2023), particularly among poor, vulnerable households (Dillon, Mueller, & Salau, 2011, Gray & Mueller, 2012a, 2012b; Dallmann & Millock, 2017; Colmer, 2021). Clement et al. (2021) estimate that climate change could force the internal migration of more than 200 million people by 2050. People similarly migrate in response to conflict, political events, and violence (Cantor & Apollo, 2020; Diepart, & Ngin, 2020; Institute for Economics and Peace, 2021). Armed conflicts may further trigger the relocation of entire households or communities (Hoogeveen, Rossi, & Sansone, 2019; UNHCR, 2019; Barrutia Barreto et al., 2022).

Importantly, internal migration is often regarded as an intermediate step to external migration — where individuals relocate internally before leaving their countries (Cattaneo & Robinson, 2019; McAuliffe & Triandafyllidou, 2021). International migration can then be considered an extension of domestic migration, driven by similar motivating factors (Lozano Ascencio, 2004; Raymer & Baffour, 2018). At the aggregate level, however, internal and external migration can complement or substitute one another (Bernard & Perales, 2021).

Understanding migration decisions is complex, as it is ultimately a multidimensional phenomenon related to a wide range of factors, including push factors —encouraging people to move out of their current location— and pull factors —attracting people to move into a new location— (Rubenstein, 2017). These factors are generally grouped into four categories: economic (e.g., job opportunities, wages); environmental (e.g., food availability, weather shocks); social (e.g., family and social networks, availability of services, quality of life); and safety/cultural (e.g., political stability, crime). Moreover, these factors are typically interrelated, vary over time, often reinforce one another, and are not always directly observable. Specific triggering factors may also be context-specific and vary by region or country (Hernandez et al., 2023). Ultimately, the decision to migrate can be regarded as a latent (hidden) decision process correlated with multiple factors, some of which push an individual to migrate.

From a policy perspective, anticipating migration decisions can be useful for targeting development programs and interventions or as a monitoring tool, to capture the migration 'sentiment' among specific groups of interest. However, making direct questions about migration intentions can be prone to refusals or underreporting. A comprehensive, holistic approach is thus needed in this regard, one that considers all potential factors at the individual, household, local, and national level driving migration decisions (Ceballos & Hernandez, 2020).

This paper describes the construction of an internal Migration Propensity Index (hereafter i-MPI) for Guatemala and Honduras. The i-MPI aims to objectively measure and track the probability of individuals in a household of permanently migrating to a different location within their country, based on a small subset of household indicators and conditions that highly correlate with the (latent) decision to migrate domestically and, combined, best predict internal migration at the household level. The methodology to develop and calibrate the i-MPI builds on the earlier construction of a MPI to predict external or cross-border migration in both Guatemala and Honduras (see Ceballos & Hernandez, 2020; and Almanzar et al., 2022), tools that have already been piloted and implemented, and are currently being validated across field surveys in the two countries.

As the original MPI for external migration, the i-MPI aims to avoid standard questions about migration intentions, which may be culturally sensitive, prone to misreporting, and difficult to monitor over time. Alternatively, the i-MPI relies on a concise, reliable, and statistically-robust scoring index obtained from indirect indicators highly associated with migration decisions that are simple and easy to collect in the field. Based on the estimated scores, a household can be labeled as being of a low, medium, or high risk of migration. All in all, the i-MPI approximates the likelihood of internal migration at the household level and can be a convenient and powerful tool for both targeting and monitoring purposes through small- and large-scale household surveys. Considering that the intention to migrate can be regarded as a proxy for local economic opportunities relative to relocating, the tool can guide investment prioritization for policy makers, donors, and program implementors on the ground, by providing a rapid indication of a household's propensity to migrate that can be further monitored over time.

The remainder of the note is organized as follows. Section 2 discusses the methodology followed to derive the index, including the statistical methods and validation procedures used. Section 3 describes the data used to calibrate the i-MPI for Guatemala and Honduras, while Section 4 presents the results of the i-MPI model for both countries. Section 5 outlines the steps to implement the index for monitoring and targeting purposes and provides a comparative analysis of the scores and indicators of migration. Section 6 offers some concluding remarks.

2. Methodology

We follow the overall approach developed for the original MPI (which was designed to predict cross-border migration). Here, we outline the basic steps for implementing said methodology but leave out the details, which an interested reader can find elsewhere (see Ceballos & Hernandez, 2020; Almanzar et al., 2022).

The general goal is to derive a simple, concise, and statistically robust index —based on easily measurable, reliable household characteristics and conditions— that can predict with reasonable confidence whether an individual in a household will migrate domestically in the near future. The index must be *simple*, to allow for donors, policy makers, and program implementors to understand and apply it for monitoring and/or targeting in real time. In addition, the index must be *concise*, to enable low-cost data collection in remote, rural areas, avoiding having to implement costly, long household surveys in

person. In this sense, the MPI comprises only ten variables carefully selected to maximize its predictive capacity based on currently available data. These variables must be *easy to measure* through direct, non-invasive questions and with minimal reporting bias. Finally, the index must be *statistically robust*, considering model fit and likely confounding effects (endogeneity). In this regard, the MPI circumvents endogeneity problems by relying on a large, detailed panel survey of households from key regions of interest, which permits to compare a household's *ex ante* characteristics with their *ex-post* decision to migrate. Furthermore, to avoid potential issues around model overfitting, the MPI is constructed and evaluated using cross-validation techniques, which split the working sample into two subsamples: one for model fitting and the other for evaluating the model's predictive capacity.

The MPI methodology departs from a standard logistical regression framework (Logit model). Since the propensity to migrate may vary according to local economic, social, and infrastructure conditions, region fixed effects (in the case of Guatemala) and department fixed effects (in the case of Honduras) are included. In particular, in the case of Guatemala we include a separate dummy for each of the following four regions: "Centro and Pacífico", "Altiplano Occidental", "Corredor Seco and Izabal", and "Verapaces".[1] In the case of Honduras, the dummies are defined across the following six departments: Copán, Intibucá, La Paz, Lempira, Ocotepeque, Santa Bárbara. All specifications also include an indicator variable to distinguish rural from urban areas, in an identical way to the external MPI in each respective country. Finally, since some households are included twice in the working sample (i.e., when a household was interviewed across all three survey rounds), we cluster standard errors at the household level to account for the likely correlation in observed outcomes within the same household.

We rely on an out-of-sample cross-validation procedure to select the final model. This is a standard statistical approach to assess the predictive performance of a model. The idea is to assess how the model performs when applied to new data. To achieve this, the data is randomly partitioned into two subsamples: 70% of the observations are used for estimating the model (i.e., model training); and the resulting coefficients are applied to the explanatory variables in the remaining 30% of the sample to test whether the model predictions match with the actual observed outcomes (i.e., model predictive performance). To minimize sampling bias, the procedure of splitting the sample is repeated 300 times.[2] In each of these repetitions, the coefficients of the model are estimated with 70% of the observations and the metrics on the model's predictive accuracy are calculated over the remaining 30% of observations. For each metric, we average the measures obtained across the 300 repetitions to derive an overall metric. We rely on the concordance statistic (C-Stat) to evaluate model fit, a metric commonly used for binary-choice models since it measures the concordance between the observed outcome and the model's predicted probability for that outcome across all observations. Models that maximize this metric exhibit the most internally-coherent array of predicted probabilities, in the sense that observations with a positive outcome (households having a migrant in our case) have on average a higher predicted probability than those with a negative outcome.[3]

---

[1] Given that the panel survey we rely on does not cover the entire country and was not representative at the department level, we choose to include fixed effects at the regional instead of the department level. Some regions are composite regions, in cases where the number of observations for a department was too small to consider it as a separate region. The description of the panel survey is presented in Section 3.

[2] These partitions naturally maintain the full-sample proportions of households with and without a migrant.

[3] The C-Stat ranges from 0.5 to 1. See Hosmer and Lemeshow (2000) for further details on the definition and interpretation of the C-Stat.

In order to maximize the objectiveness and unbiasedness of the final selected model, we rely on a stepwise statistical procedure to select the variables with the highest predictive power from amongst the universe of variables considered in each setting. In particular, we implement the following steps:

1. Consider every possible one-variable model, constructed by only including one single variable from the universe, in addition to the location fixed effects and the rural indicator.
2. Estimate each of these one-variable models and calculate their out-of-sample predictive-power metric.
3. Select the one-variable model with the highest predictive power.
4. Add one additional variable (from the remaining variable universe) to the selected one-variable model, to construct all possible two-variable models.
5. Estimate each of these two-variable models and calculate their out-of-sample predictive power metric.
6. Select the two-variable model with the highest predictive power.
7. Add one additional variable to the selected two-variable model (from the remaining variable universe), to construct all possible three-variable models and repeat the selection process, adding one variable at a time and selecting the one with the highest predictive power in each round, until a 10-variable model is achieved.

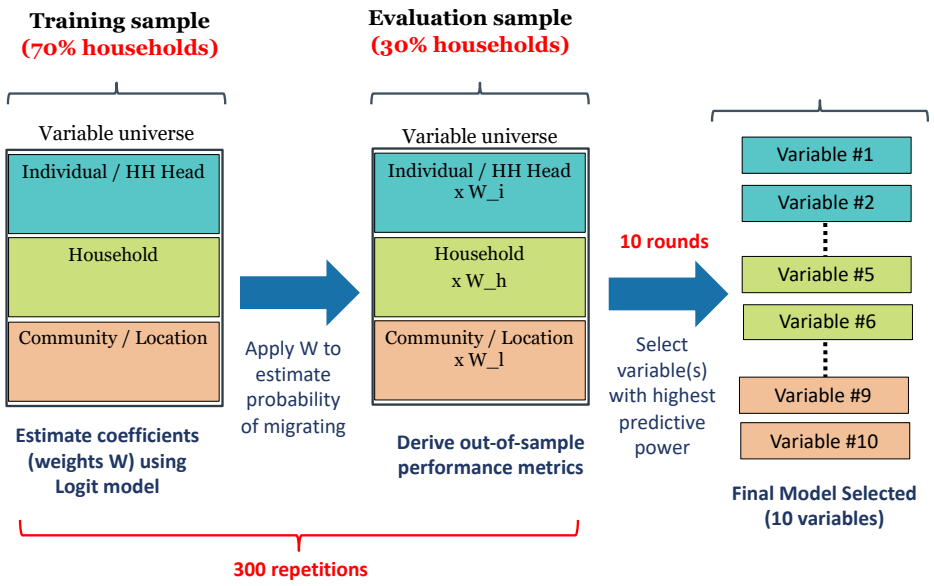
The above steps provide an objective procedure on which to build a statistically-robust predictive model. However, in order to fulfill all index goals, we allow for a degree of subjectivity in the process. Besides considering variables that are simple and easy to measure, we aim to obtain variables that are also (1) relatively dynamic over time, (2) add value to the model, and (3) cover key dimensions around the decision to migrate.

In terms of the first element, since one of the objectives of the index is to monitor the propensity of an individual in a household to migrate over time, it is important that at least a subset of the selected variables likely changes over time. For instance, the sex of the household head or his/her education would normally not change from one year to the next, while ownership of certain assets or exposure to natural shocks may. As for adding value to the model, we intend to avoid situations where two similar indicators (likely highly correlated) are jointly included. For instance, if the model already includes a variable capturing whether any of the household members in a certain age group has completed 5 or more years of education, additionally including whether any member has completed 6 or more years of education would not add real value. Finally, we attempt to capture most key dimensions around the decision to migrate in the indicators proposed, including factors generally discussed in the broader migration literature and in studies specific to the Northern Triangle.

To achieve this, we follow the statistical procedure described to select the optimal variable at each round but we further assess the new selected variable in light of the objectives outlined above; if the new variable diverts from any of these objectives it is then excluded, and the procedure is allowed to select the second-best variable for that round.[4] This process permits to have a final set of 10 variables that fulfill all the i-MPI objectives. Figure 1 summarizes the complete methodology.

[4] The whole estimation process is documented and is available upon request.

Figure 1: Description of the internal Migration Propensity Index (i-MPI) procedure



3. Data

This section describes the data sources on which we draw to calibrate the i-MPI in both countries. As discussed above, for the MPI calibrations we rely on longitudinal panel data, where the same set of households is observed during at least two points in time. The panel nature of the data allows to account for the timing of migration decisions when identifying the factors correlated with these, avoiding potential confounding effects stemming from changes in household characteristics occurring post migration. For instance, when an individual in a household migrates, the household structure and dynamics may immediately change, and the household may begin to receive remittances shortly after. Under this scenario, panel data provides the ability to observe migration instances in the recent past (captured in a later survey) in addition to household characteristics and conditions prior to the migration event (captured in an earlier survey), allowing for a better identification of the drivers triggering the migration decision at the time it was made.

Once the data source for the calibration is selected, the next step is to construct the so-called 'variable universe', which encompasses all available variables in the survey that could potentially be correlated with the decision to migrate by an individual and that should be relatively easy to collect in future surveys. These variables are normally related to common migration drivers, including economic, environmental, social, safety, and cultural factors. While the variable universe for each country is exhaustive of all available indicators likely associated with the decision to migrate, some important factors may be omitted since the available surveys were geared towards other development objectives and not designed to monitor migration per se. In this regard, the resulting i-MPI tool (along with the earlier versions considering external migration), should be viewed as an initial attempt to construct a robust predictive model, which can be further refined over time as panel data with a wider set of indicators become available.

While many of the indicators considered in the universe are related to local push factors, there is a small subset of variables aimed at capturing certain pull factors. For instance, we consider whether the household knows anyone in their community that migrated abroad or domestically during the previous year or whether other households in the community receive remittances from abroad, as well as whether

the household itself receives remittances. These variables can, for instance, capture the effect of social ties with individuals from the community who have successfully migrated domestically or abroad (i.e. peer or network effects), which is known to be an important pull factor prompting other individuals in the community to follow suit.[5] In terms of broader pull factors, such as policies or job opportunities in other regions or abroad, we argue that, while important, they generally exert a blanket effect on attracting migrants, affecting all households in a similar manner and thus acting as an overall (positive or negative) shifter to the probability to migrate. As such, the MPI remains a valuable tool for identifying those households most prone to emigrating within a given group, independent of the current external conditions.

Finally, we discuss the construction of the domestic migration indicator for both settings, which is used as the dependent variable to calibrate the statistical models and derive their predictive power. In the case of Guatemala, the outcome relates to migration instances during the 12 months prior to the date of the interview, while in the case of Honduras it relates to a period of either 12 or 24 months, depending on the survey wave. For this reason, the methodology for Guatemala estimates the probability that an individual in a household emigrates during the following 12 months, while the one in Honduras is predictive of the decision to migrate within the following 24 months. Henceforth, we thus refer to the calibrated i-MPI more generally as capturing the decision to migrate in the near future.

Guatemala

In the case of Guatemala, we work with three survey rounds collected among the same group of households between 2012 and 2014, as part of the impact evaluation of the Plan Pacto Hambre Cero (PPH0).[6] The survey was designed by IFPRI in coordination with the Secretary of Food Security and Nutrition (SESAN) and implemented by the National Statistical Institute (INE). While the PPH0 focused on the 176 municipalities with the highest stunting rates in the country, the panel dataset includes a subsample of 60 municipalities where a specific number of households were followed over time.[7] The targeted population of the PPH0 were households that have children under five years old, pregnant women, and/or women between 15 to 49 years old, and the final panel dataset is representative of this subpopulation within the selected municipalities.[8] More details on the Guatemala dataset are described in Ceballos & Hernandez (2020).

The surveys were implemented during November and December 2012, 2013, and 2014, and included modules on household socioeconomic characteristics, income sources, expenditures and assets, dwelling characteristics, access to services and credit, participation in social programs, hygiene and

[5] See, for example, Cohn et al. (2017), Congressional Research Service (2019a, 2019b), Creative Associates International (2019), and Clemens (2021).

[6] The PPH0 was a national strategy designed by the Government at the beginning of 2012 to address the problem of chronic malnutrition, acute malnutrition, and food insecurity with special emphasis on households with children under five years old, pregnant women and women of childbearing age (between 15 to 49 years old).

[7] Besides food insecurity and malnutrition, these 176 municipalities also fall behind in other development indicators such as prevalence of poverty, access to basic services, or market accessibility, reason why we refer to them as “vulnerable” municipalities. For monitoring purposes of the PPH0, each survey round also covered other municipalities that were part of the 176 targeted municipalities, however, these were randomly selected every year and were not followed over time.

[8] In practice, this is close to the entire household population given that almost 89% of households in the listing had had at least one member in these targeted groups.

family health, nutrition, and food consumption. The final panel dataset includes 1,344 households interviewed during both 2012 and 2013 and 1,454 households interviewed during both 2013 and 2014, for a total of 2,798 households (with 779 households interviewed during all three years).

The variable universe for Guatemala includes a total of 57 variables related to household socioeconomic characteristics and composition, dwelling materials, access to services, assets, landholdings and agricultural activities, expenditures, participation in social programs, climate vulnerability, whether someone emigrated during the past year, incoming remittances, and geographic location. The final variable list, together with each variable's definition and type of measure, is presented in Appendix A.

To capture whether a household had an instance of domestic migration, we rely on the following two questions, asked in each survey round:

- Did any household member leave the community to seek work either temporarily or permanently during the last 12 months? *Yes / No*
- *[If "Yes"]* Where did they go? *(A community in…)*
    1. *Same municipality*
    2. *Same department*
    3. *Another department*
    4. *Guatemala City*
    5. *Another country*

We define a household as having had an internal migrant during the previous 12 months if the answers to the first question above is "Yes" and the answer to the second question is either (another community in…) 'Another department' or 'Guatemala City'. We purposely do not include cases of migration to another community in the same municipality or in the same department, since such situations may include relocations related to marriages or local job opportunities which are outside the scope of the type of domestic migration of interest.

Honduras

In the case of Honduras, we rely on a panel dataset collected by IFPRI for the impact evaluation of USAID-ACCESO across three survey rounds in 2012, 2013, and 2015. Henceforth referred to as IFPRI-FTF, the surveys encompassed six departments in western Honduras where the activities under Feed the Future/USAID were taking place, i.e., Copán, Intibucá, La Paz, Lempira, Ocotepeque, and Santa Bárbara, and were designed to be representative at the department level. More details on the Honduras dataset are described in Almanzar et al., (2022).

The surveys were collected during Honduras' lean agricultural season, spanning May through August each year. These included modules on household socioeconomic characteristics, income sources, occupations, expenditures, assets, dwelling characteristics, access to services and credit, participation in social programs, hygiene and family health, nutrition, and food consumption, among others. The final panel dataset comprises 2,496 households for which data is available across all three survey waves.

Another 327 additional households are present in just the initial two survey waves, 2012 and 2013, for a total of 2,823 in the final panel sample.

The variable universe for Honduras includes a total of 64 variables related to household socioeconomic characteristics and composition, dwelling materials, access to services, assets, landholdings and agricultural activities, expenditures, participation in social programs, climate vulnerability, whether someone migrated domestically or internationally during the past year, incoming remittances, and geographic location. The final variable list, together with each variable's definition and type of measure, is presented in Appendix B.

To capture whether a household had an instance of domestic migration, we rely on differences between household members' lists between two consecutive survey rounds and on a follow-up question inquiring about the reason for which an individual that used to be a member of the household is no longer present, as follows:

- Why is [previous member] no longer a household member?
    1. *He/she passed away*
    2. *He/she was never a member; was in the roster by mistake*
    3. *He/she moved to another HH inside this village*
    4. *He/she moved to a different municipality*
    5. *He/she moved to Tegucigalpa*
    6. *He/she moved to a different department*
    7. *He/she moved to a different country (specify the country)*

We define a household as having had an internal migrant since the previous survey round if there is at least one individual that belonged to the household during the previous survey round and is no longer a member of the household in this survey wave, and the answer to the question above was (5) he/she moved to Tegucigalpa, or (6) he/she moved to a different department. Similarly to the case of Guatemala, we exclude other cases of migration to the same village or another community in the same or another municipality.

4. Results

This section presents the results of the i-MPI calibration exercise outlined above. Table 1 lists the final 10 indicators selected for the i-MPI Guatemala (panel A) and Honduras (panel B). The variables are shown in the order in which they were selected, based on the C-Stat metric (which maximizes the model's internal coherence). The table also indicates whether the effect of that variable on the probability to migrate domestically is positive or negative. These 10 variables, together with the rural indicator and department shifters, conform the final model in each country.

Table 1: Indicators selected for the internal Migration Propensity Index (i-MPI)

Panel A. Guatemala

#	Indicator	Effect on probability to migrate
1	Community-level rate of domestic (within country) migration	**Positive**
2	Household has received remittances in the last 12 months	**Negative**
3	Household owns a cellphone	**Positive**
4	Dwelling has finished walls (block or brick)	**Negative**
5	Number of household members	**Positive**
6	At least one household member aged 30-49	**Negative**
7	At least one female household member has one year of secondary education	**Positive**
8	Household is beneficiary of social program "Bono Seguro de Salud"	**Positive**
9	Majority of households in the community connected to sewerage	**Negative**
10	At least one male household member has no education	**Negative**
+	**Rural and region indicators**	**(Shifters)**

Panel B. Honduras

#	Indicator	Effect on probability to migrate
1	Number of household members	**Positive**
2	At least one household member approved one year of secondary education	**Positive**
3	Age of household head in years	**Positive**
	Square of age of household head in years	**Negative**
4	Household experienced a drought event on previous primera ag. season (SPI<-1)	**Positive**
5	Household owns a vehicle (car/motorcycle)	**Negative**
6	At least one household member aged 15-29	**Positive**
7	At least one household member has a permanent disability	**Negative**
8	Household owns the house in which they live	**Negative**
9	Household owns a cellphone	**Positive**
10	Number of completed years of education of household head	**Positive**
	Square of number of completed years of education of household head	**Negative**
+	**Rural and department indicators**	**(Shifters)**

Figure 2 shows, in turn, the marginal effect of the selected variables on the probability to migrate domestically in the final model specification for Guatemala (panel A) and Honduras (panel B), as estimated using the full sample. The marginal effects represent the estimated change in the outcome variable stemming from a change in an explanatory variable, with the other variables held constant at their respective means. It is important to note that, since the models are ultimately calibrated on

observational data, these marginal effects are not to be interpreted as causal estimates of the effect of the selected variables on the probability to migrate.

In the case of Guatemala (panel A), variables capturing peer effects or the influence of a household's network (pull factors) stand out as important predictors of the probability that someone migrates to another department or to the capital city in the near future. The community-level rate of internal migration is strongly and positively associated with a household's probability to migrate domestically, and having a migrant abroad (proxied by whether the household received remittances during the last 12 months) is associated with a lower probability to do so (though this same variable is associated with a higher probability to migrate abroad, as captured in the external MPI for Guatemala). Certain demographic characteristics also turn out to be key factors: a larger household size, having a female member that has at least one year of secondary education, having no male members with no education, or having no members in the 30-49 age group, are all associated with a larger propensity to migrate internally. This is consistent with much of the existing literature on migration, which finds youth as the main age group migrating, usually with a moderate level of education and skills. Owning a cellphone is also associated positively with the propensity to migrate, in addition to being the recipient of a government social program. Finally, having finished walls in the dwelling (as captured by the walls' material being block or brick) and living in a community where the majority of households are connected to sewerage are negatively associated with the probability of a family member migrating domestically. The latter four variables can be generally interpreted as proxies that jointly capture a household's socioeconomic level.

In the case of Honduras (panel B), pull factors were not selected as being predictive of internal migration, despite their inclusion in the variable universe.[9] Nonetheless, some of the other characteristics are similar to those in the i-MPI Guatemala. Regarding family composition and other demographic characteristics, larger households, those with at least one member between 15-29 years of age, and those with at least one member having at least one year of secondary education are all associated with a larger propensity to migrate domestically. A distinctive feature worth mentioning is that in Honduras the calibration procedure selected the head of household's age and years of education both in levels and squared, indicating the presence of non-linear effects of these variables on the probability to migrate. In particular, a positive effect in levels and a negative effect for the square indicates a positive association with the probability to migrate which gradually diminishes over the range of values taken by each of these two indicators. The presence of a household member with a permanent disability or chronic condition (e.g., visual, hearing or mobility impairment, loss of limbs, or chronic illness) is negatively associated with a household's propensity to migrate. Similarly to Guatemala, having a cellphone is positively associated with the probability to migrate domestically, while ownership of a motorized vehicle (car or motorcycle) and owning the house where the household lives both reduce the likelihood that an individual in that household migrates. Finally, experiencing a mild drought event in the previous *primera* agricultural season, as captured by the standardized precipitation index being smaller than -1 (i.e., one standard deviation below the historical average), is associated to a higher probability to migrate internally, with one

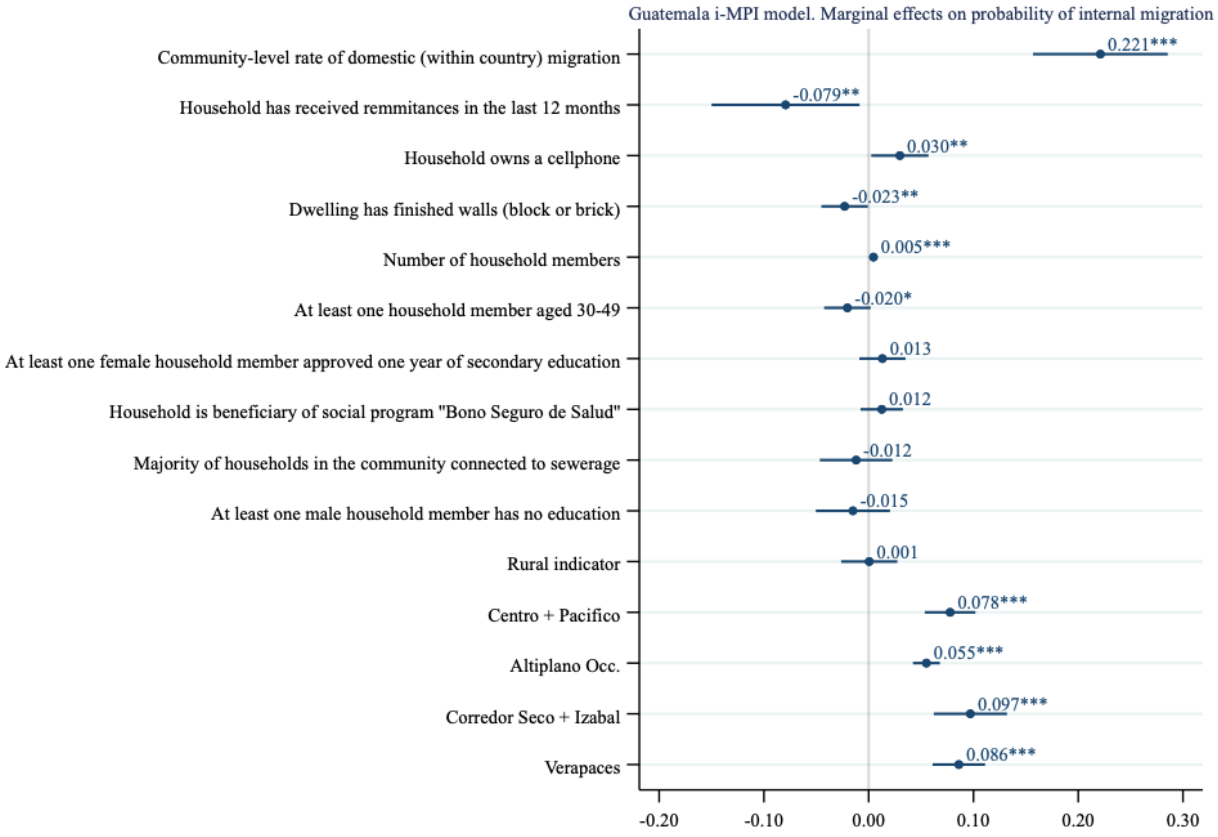
[9] The variable universe included the community-level share of households that had reported an internal migrant and whether the household itself had received remittances in the 12 months prior to the interview, but these variables were not selected among the ten most relevant by the calibration procedure. The construction of community-level variables in Honduras was constrained by the geographic definition available in the surveys, where we rely on a broader group-unit ('Aldea', with an average of 9 and a standard deviation of 15 households) than in Guatemala (group-unit of 'Sector', with an average of 8 and a standard deviation of 3 households). This broader definition of community could hinder the informativeness of these community-level indicators.

of the largest marginal effects among the variables chosen, highlighting the importance of adverse climatic events in the decision to migrate.[10]

Table 2 shows performance indicators for the final i-MPI Guatemala (panel A) and i-MPI Honduras (panel B) models. Each panel first reports the full in-sample metrics; that is, the metrics using coefficients from a model calibrated with the full working sample (without considering the cross-validation splits). Next, the in-sample (training) and out-of-sample metrics are reported, using the 70%-30% sample partitions for the cross-validation procedure; with the average and confidence interval across the 300 replicated subsamples. Finally, each panel reports 'Sensitivity', 'Specificity', and 'Correct Classification Rate' measures of the final models.[11]

Figure 2. Marginal effects statistics of the i-MPI model for Guatemala and Honduras

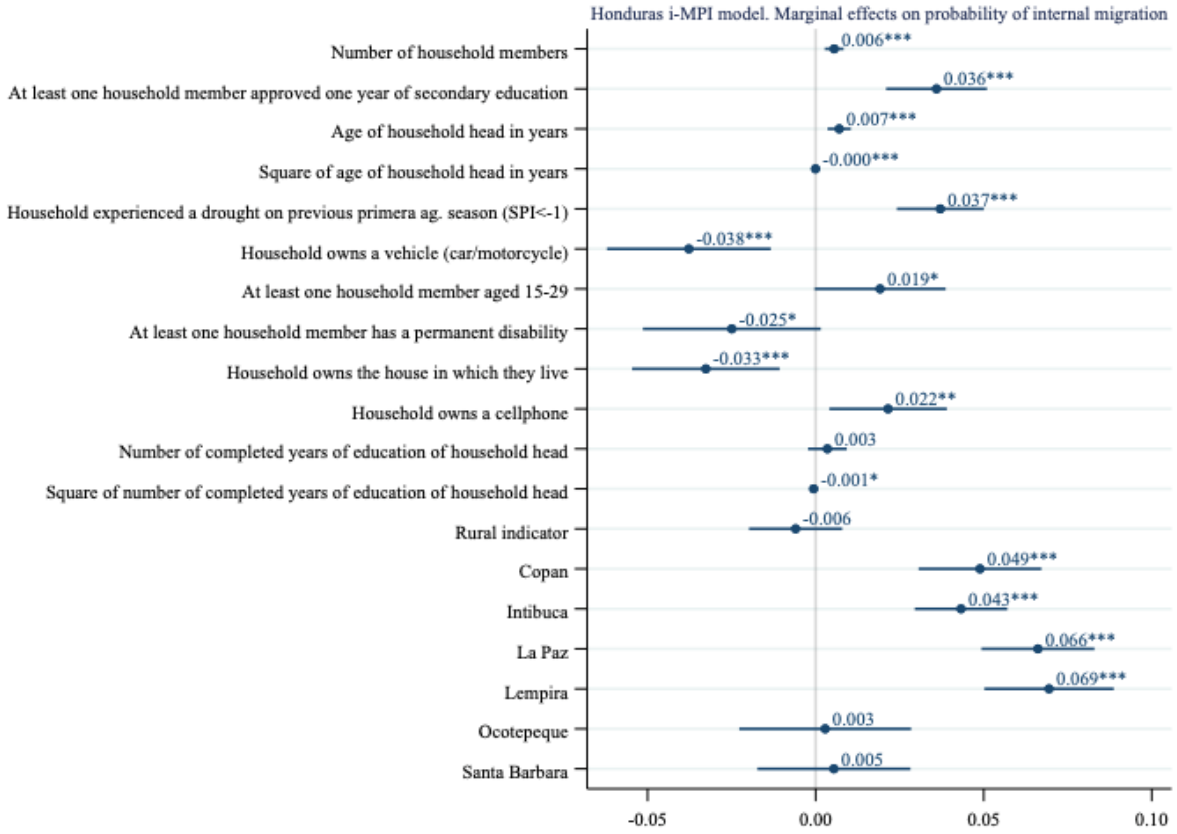
Panel A. Guatemala



[10] This latter variable was also selected as an important predictor of external migration in the MPI Honduras (see Almanzar et al., 2022).

[11] These indicators crucially depend on deriving a binary model prediction for each household to compare with observed migration instances, for which we use a value of 0.1, or the probability threshold above which the i-MPI considers a household to be of medium risk of migrating, equivalent to 50 points.

Panel B. Honduras



Note: Panels A and B plot marginal effects of the selected variables on the probability to migrate domestically in the final model specification, based on Logit regression coefficients for each country. The bars indicate 95% confidence bounds based on standard errors clustered at household level. The estimation sample comprises 2,702 observations in Guatemala (panel A) and 5,016 observations in Honduras (panel B). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

Overall, the final models show a good predictive performance. In particular, the models' sensitivity (or their ability to correctly predict positive cases, i.e. households in the sample where an internal migration instance is observed) is in the order of 43% in Guatemala and 44% in Honduras. In other words, the model for each country is able, by tracking the 10 variables alone, to *ex ante* correctly identify more than two in every five households where someone ends up migrating.[12] At the same time, the models' specificity (or their ability to correctly predict negative cases, i.e. households whose members did not migrate) is 83% in Guatemala and over 85% in Honduras. That is, the model correctly identifies a non-migrating household in more than four out of five cases. It is important to clarify that when the prevalence of the modeled outcome is low, the amount of false-positives tends to be larger, while the number of false-negative outcomes is lower. In this sense, a model with a 44% sensitivity is a fairly accurate one in relative terms. Ultimately, both the sensitivity and the specificity measures are important for policy and targeting purposes: correctly identifying households with future migrants is relevant for accurately targeting programs aimed at reducing incentives for migration, while correctly identifying non-migrating

[12] We obtain similar sensitivity rates in the external MPI calibration for Guatemala and Honduras.

households is pertinent to better allocate program resources, which can be used in more efficient ways or to achieve other development objectives.

Table 2. In-sample and out-of-sample performance measures of selected models

Panel A. Guatemala

	In-sample	Training In-sample		Out-of-sample	
	Full Sample	Mean	95% conf. interval	Mean	95% conf. interval
Root Mean Square Error	0.249	0.249	[0.239-0.261]	0.250	[0.223-0.272]
Concordance Stat.	0.728	0.730	[0.709-0.755]	0.700	[0.638-0.766]
Sensitivity	0.428	0.440	[0.371-0.510]	0.396	[0.269-0.538]
Specificity	0.835	0.830	[0.790-0.861]	0.828	[0.776-0.875]
Correct Classification Rate	0.806	0.803	[0.767-0.829]	0.798	[0.757-0.837]
Number of observations	2,702				
Number of households	1,955				
Number of internal migrants	191				

Panel B. Honduras

	In-sample	Training In-sample		Out-of-sample	
	Full Sample	Mean	95% conf. interval	Mean	95% conf. interval
Root Mean Square Error	0.222	0.222	[0.214-0.230]	0.223	[0.203-0.241]
Concordance Stat.	0.755	0.757	[0.739-0.775]	0.738	[0.693-0.783]
Sensitivity	0.438	0.459	[0.401-0.513]	0.424	[0.334-0.518]
Specificity	0.859	0.858	[0.838-0.879]	0.857	[0.832-0.884]
Correct Classification Rate	0.836	0.836	[0.817-0.856]	0.833	[0.812-0.855]
Number of observations	5,016				
Number of households	2,631				
Number of internal migrants	281				

Note: The sensitivity (correct positive outcome rate), specificity (correct negative outcome rate), and correct classification rate (correct positive and negative outcomes rate) are calculated using a threshold of 0.1 to convert the estimated model probability into a binary prediction and compare them with the observed (actual) outcomes.

5. Scoring

This section presents the final i-MPI questionnaires and the scoring rules to calculate the likelihood of a household migrating domestically in the near future. Panels A and B of Figure 3 show the i-MPI questionnaire for, respectively, Guatemala and Honduras, to be asked to any given household for which there is interest in determining their propensity to migrate internally. The questions are simple, non-invasive, and easy to collect, without direct inquiries about intentions or attempts to migrate. The questionnaire can be easily implemented either in person or over the phone.

Figure 3. Internal Migration Propensity Index (i-MPI) questionnaires

Panel A. Guatemala

	Question	Answer	Instructions	Score
a	Approximately how many households live in your community?	___ households		
b	In how many of these households did someone migrate to another Department in Guatemala or to Guatemala City during the last 12 months?	___ households	Calculate: (b / a) x 30	
c	How many people live in this household in total?	___ members	Calculate: c x 0.6	
d	Has any member of your household received remittances in the last 12 months?	No	11	
		Yes	0	
e	Does any household member have between 30 and 49 years of age?	No	3	
		Yes	0	
f	Of all the WOMEN over 15 years of age in this household, has anyone completed at least the first grade of *secondary/high school* education?	No	0	
		Yes	2	
g	Of all the MEN over 15 years of age in this household, has anyone not completed at least the first grade of *primary* education??	No	2	
		Yes	0	
h	Is your household a beneficiary of Bolsa Segura or another social program from the Guatemalan government?	No	0	
		Yes	2	
i	Are the exterior walls of the house made of BRICK/BLOCK?	No	3	
		Yes	0	
j	Does any household member own a cellular phone?	No	0	
		Yes	4	
k	Are the majority of households in your community connected to sewerage?	No	2	
		Yes	0	
l	Is the household located in a rural area?	No	0	
		Yes	0.1	
m	In which region is this household located?	Centro + Pacifico	38	
		Verapaces	39	
		Corredor Seco + Izabal	40	
		Altiplano Occidental	35	

Under 50	**Low risk**	Between 50 and 70	**Medium risk**	Above 70	**High risk**	**Total score:**	

Panel B. Honduras

	Question	Answer	Instructions	Score
a	How many people live in this household in total?	___ people	Calculate: a x 0.9	
b	Does any household member own the house where you live?	No	5	
		Yes	0	
c	Does any household member have between 15 and 29 years of age?	No	0	
		Yes	3	
d	Of all household members over 15 years of age in this household, has anyone completed at least the first grade of secondary/high school education?	No	0	
		Yes	6	
e	What is the age of the head of household (in years)?	___ years		
		Calculate: 1.2 x age - 0.01 x [age x age]		
f	How many years of education does the head of the household have?	___ completed years		
		Calculate: 0.6 x years - 0.1 x [years x years]		
g	Does any household member suffer from a permanent disability? (visual, hearing or mobility impairment, loss of limbs, or chronic illness)	No	4	
		Yes	0	
h	Does any household member own a car or motorcycle?	No	6	
		Yes	0	
i	Does any household member own a cellular phone?	No	0	
		Yes	4	
j	In the last *primera* agricultural season (May to September) was there a month in which it rained considerably less than usual?	No	0	
		Yes	6	
k	Is the household located in a rural area?	No	1	
		Yes	0	
l	In which region is this household located?	Copán	3	
		Intibucá	2	
		La Paz	6	
		Lempira	6	
		Ocotepeque	3	
		Santa Bárbara	4	

Under 50	**Low risk**	Between 50 and 70	**Medium risk**	Above 70	**High riskue**	**Total score**	

The scores for each question are derived directly from the estimated statistical model. To do this, the full set of logit coefficients is re-scaled to obtain an i-MPI score that aligns with the underlying probability thresholds from the logistic model.[13] In particular, an i-MPI score of 50 —below which households are considered to be of low risk to migrate domestically—, is equivalent to a probability of 0.01 in the logit model. At the same time, an i-MPI score of 70 —above which households are considered to be of a high internal migration risk—, correspond to a logit probability of 0.1. These thresholds are the same for both the i-MPI Guatemala and the i-MPI Honduras, and strike a balance between optimizing the model's performance at correctly identifying households where an individual migrates in the near future (specificity) and correctly identifying those where an individual does not migrate (sensitivity). In addition, they represent natural breaks in the effective probability of migration (estimated as the fraction of households around a given score that reported an effective domestic migrant) among households in the sample. To better appreciate these breaks, Figure 4 depicts local polynomial plots between the effective probability of migration within the country and i-MPI scores.[14] The total score normally ranges from 0 to 100, where 100 represents a higher propensity or likelihood to migrate.[15]

After obtaining the answer to each question in the i-MPI questionnaire, the scoring instructions to the right of each question must be followed to obtain the partial score contributed by that question. Once all responses are collected or derived, the total score should be added up and, by using the box at the bottom of the questionnaire, the migration risk category of the household can be determined. The scores for some of the variables, such as the household's size, or household head's age or years of education must be derived through a simple formula. If performing these calculations in the field is problematic or prone to errors, it is possible to collect the answers and process them in a centralized location. An accompanying automated excel sheet is provided for these purposes.

It is important to mention that the questions around the number of households in the community and the number of these households with a recent migrant are meant to provide rough estimates for the proportion of households in a community with a recent migrant, avoiding the need to interview all households in a community. Now, when the community under consideration is fairly large, this question could focus on a narrower definition of the community, such as those households with whom the interviewee normally interacts with (as this variable is mainly intended to proxy for potential peer or network effects in migration decisions); or could be replaced altogether by community-level estimates of the corresponding proportion, in case these data were available. Similarly, the question on whether there was a month in the previous *primera* season when it rained considerably less than usual could be replaced by an indicator of whether the standardized precipitation index (calculated from external sources) for any of these months was below -1. Such approaches could help reduce some of the measurement error in the calculation of the index.

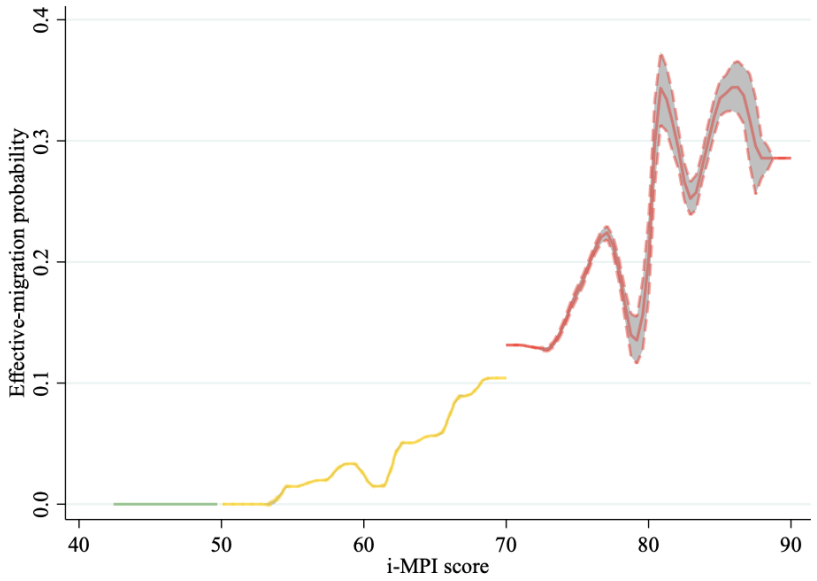
[13] Some of the final i-MPI scores are not exactly equivalent to the ones in the original model since some minor rounding is done to simplify and avoid multiple decimals in the scoring system.

[14] The local polynomial plot is derived using an Epanechnikov kernel, standard rule-of-thumb bandwidth, and assuming a polynomial of degree 0.

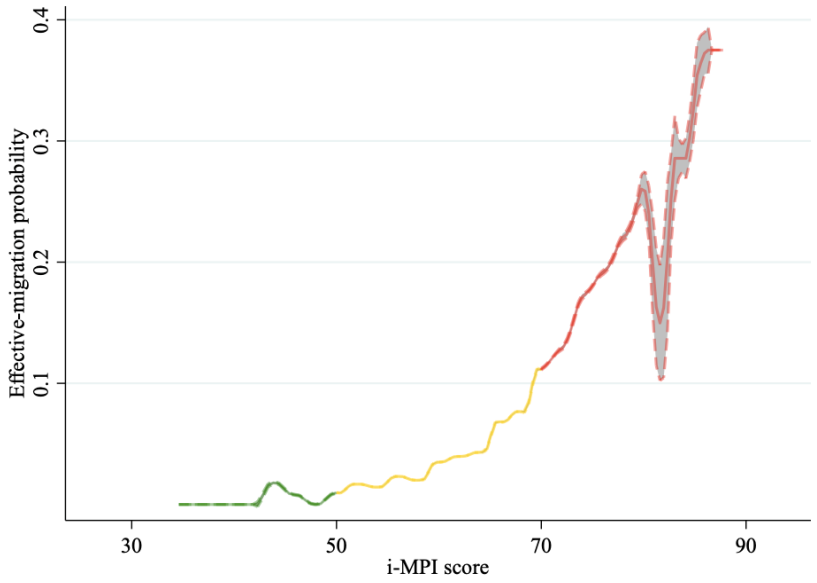
[15] The score can theoretically take on values below 0 or above 100. This would be very rare, however, since in practice the MPI for all households in the training datasets was within the 0-100 range.

Figure 4. Effective migration probability and MPI scores

Panel A. Guatemala



Panel B. Honduras



Note: This figure shows a local polynomial plot between the effective probability of migration and i-MPI scores. The local polynomial plot is derived using an Epanechnikov kernel, standard rule-of-thumb bandwidth, and assuming a polynomial of degree 0. The effective probability of migration corresponds to the rate of observed internal migrants among all households in a given range of i-MPI scores (bins). The shadowed area delimited by dashed lines represent the confidence bounds at the 95% level.

The resulting i-MPI score is directly associated to the estimated probability that at least one member within the household migrates domestically in the near future. Overall, households with i-MPI scores below 50 are considered to be of low risk in terms of their likelihood to migrate domestically. In turn, households with scores between 50 and 70 are considered of medium risk, and those with scores above 70 are considered of high internal migration risk. If desired, the estimated model probabilities can be determined through the lookup table in Appendix C for different i-MPI scores. It is important to bear in mind, however, that these probabilities are centered around the respective population-level migration rates in each country, and are thus low in absolute terms, as discussed in the previous section, so they should not be taken at face value.

If one were interested in obtaining the propensity of a given household to migrate abroad (in addition to that to migrate domestically, as estimated by the i-MPI), one can simply ask the few additional questions needed for the external MPI (not asked already in the i-MPI) and obtain its score. Nonetheless, the MPI and i-MPI scores are not directly related and are to be calculated independently. The next subsection discusses the empirical relationship between both of these scores.

Comparative analysis between i-MPI scores and internal and external migration

In this subsection, we present a descriptive and graphical analysis of the relationship between the i-MPI score, the external MPI score (hereafter, e-MPI, for the sake of exposition), and effective internal and external migration rates, for both Guatemala and Honduras. The i-MPI and MPI scores are calculated as the average of the respective household-level scores in each municipality and the effective migration rates are calculated as the percentage of households with internal or external migrants in each municipality. The purpose of these analyses is both to illustrate the i-MPI fit to the underlying internal migration data and to describe how internal and external migration relate to one another in each country.

Table 3 shows pairwise correlations between the above-mentioned indicators. Reassuringly, the i-MPI and e-MPI scores are positively correlated with, respectively, the effective rates of internal and external migration at the municipality level in both countries (panels A and B). This is not surprising, since the final set of variables in each MPI was specifically selected to best predict observed migration instances. On the other hand, the correlation between both MPI scores is negative and significant in the case of Guatemala but positive and significant in the case of Honduras.[16] In other words, while a higher propensity to migrate domestically is associated to a lower propensity to migrate internationally in Guatemala, both propensities are positively related in the case of Honduras. To an extent, this pattern reflects existing relationships in the data, where the internal and external effective migration rates are negatively and statistically-significantly associated in Guatemala, and positively (but not statistically-significantly) associated in Honduras.

[16] These relationships hold when the pairwise correlation is calculated at the household-level; we refrain from including these results here to avoid confusion between the different data analysis levels.

Table 3. Pairwise correlation matrix of municipality-level averages of i-MPI and e-MPI scores and effective internal and external migration rates

Panel A. Guatemala

	i-MPI score	e-MPI score	Effective internal migration rate	Effective external migration rate
i-MPI score	1			
e-MPI score	-0.466***	1		
Effective internal migration rate	0.775***	-0.278***	1	
Effective external migration rate	-0.253***	0.710***	-0.194**	1

Panel B. Honduras

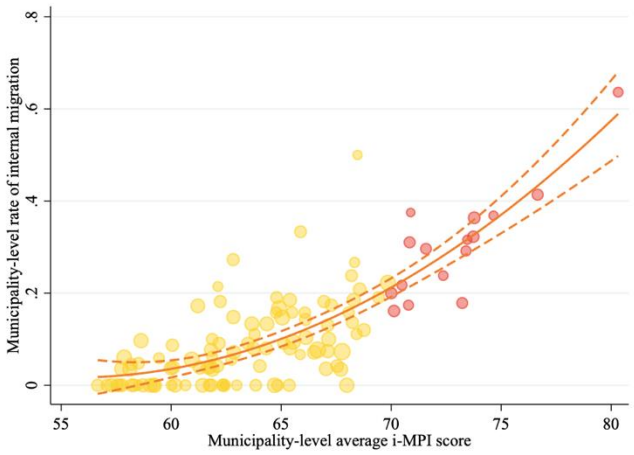
	i-MPI score	e-MPI score	Effective internal migration rate	Effective external migration rate
i-MPI score	1			
e-MPI score	0.406***	1		
Effective internal migration rate	0.361***	0.088	1	
Effective external migration rate	0.195***	0.621***	0.042	1

Note: ***, ** and * denote statistical significance at, respectively, the 1%, 5%, and 10% levels. i-MPI and e-MPI scores represent averages of these scores across all households in a given municipality. Effective migration rates are calculated as the percentage of households with internal or external migrants in a given municipality. The sample comprises 115 municipalities in Guatemala (panel A) and 174 in Honduras (panel B) with more than 5 households in each survey round.

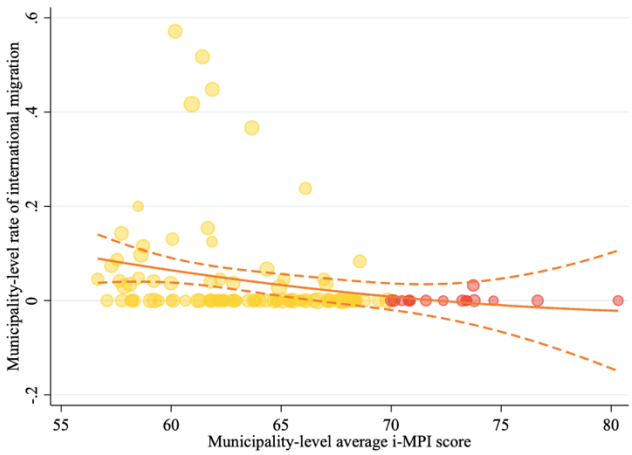
Figure 5, in turn, shows scatter plots between municipality-level rates of internal (panels A and C) and external (panels B and D) migration and municipality-level i-MPI scores in Guatemala (upper panels) and Honduras (lower panels), where each circle in the figure represents a municipality (proportional to its size). The figure includes a fitted quadratic line, together with its confidence intervals (dashed lines). As expected, the relationship between the i-MPI score is strongly associated to effective rates of internal migration in the municipality, particularly in the case of Guatemala (panel A). The figure also shows the negative relationship between the i-MPI and effective external migration rates in Guatemala, and the positive (though weaker) relationship in the case of Honduras.

Figure 5. Migration indicators and MPI scores

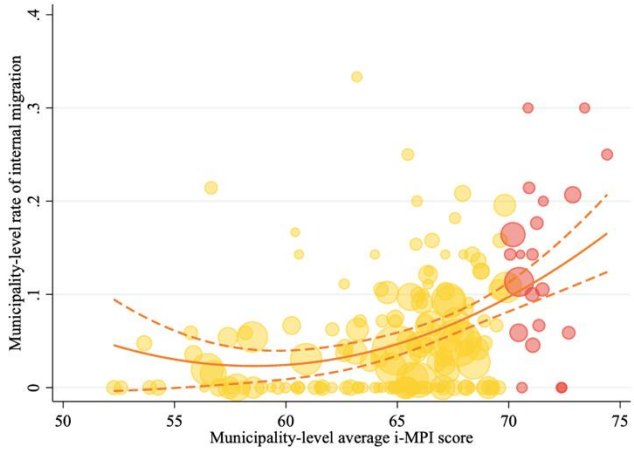
Panel A. Guatemala – i-MPI vs. Effective internal migration



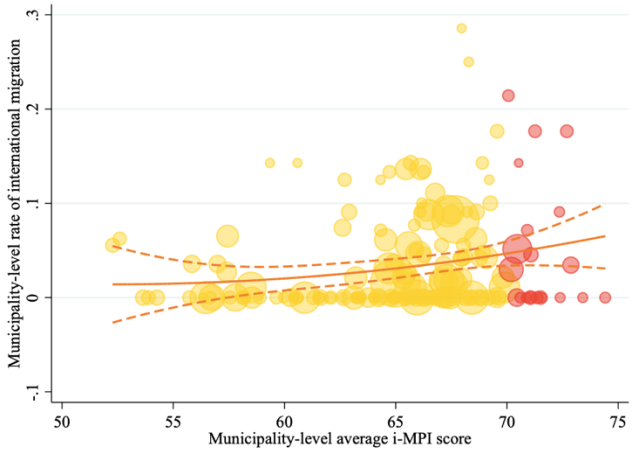
Panel B. Guatemala – i-MPI vs. Effective external migration



Panel C. Honduras – i-MPI vs. Effective internal migration



Panel D. Honduras – i-MPI vs. Effective external migration



Note: The figure shows a scatter plot and a quadratic fit line between municipality-level rates of internal (panels A and C) and external (panels B and D) migration and municipality-level averages of i-MPI scores for municipalities in Guatemala (upper panels) and Honduras (lower panels). i-MPI and e-MPI scores represent averages of these scores across all households in a given municipality. Effective migration rates are calculated as the percentage of households with internal or external migrants in a given municipality. The sample comprises 115 municipalities in Guatemala (panel A) and 174 in Honduras (panel B) with more than 5 households in each survey round. The size of the marker (circles) in the scatter plot represents the size of the municipality. The dashed lines represent 95% confidence intervals of the fitted quadratic line.

6. Concluding remarks

The Internal Migration Propensity Index (i-MPI) aims to objectively measure and track the probability of individuals in a household permanently migrating to a different location within their country, based on a small subset of household indicators and conditions that highly correlate with the (latent) decision to migrate domestically and, combined, best predict internal migration at the household level. The i-MPI shares the same objective as its counterpart for external migration, the original MPI. This note describes the design and calibration of the i-MPIs for both Guatemala and Honduras, using household panel survey datasets collected in each country.

The i-MPI is designed to provide donors, policy makers, and program implementors a convenient tool to measure and track the likelihood that individuals in a given household migrate domestically —to another department or to the capital city— over the near future. The index is based on an approach that is reliable —by avoiding sensitive, direct inquiries on migration intentions—, easy to implement —by relying on a concise and simple set of questions—, and statistically-rigorous —constructed using advanced statistical methods and evaluated in terms of its out-of-sample predictive capacity.

The i-MPI relies on ten simple, easy-to-measure questions that can be directly collected over the phone or embedded into a larger survey. Similarly, the non- invasive, indirect approach of the score allows its periodic application over time to the same population for monitoring purposes. Ultimately, the i-MPI scoring system allows to classify households as being of low, medium, or high risk of internal migration, based on the estimated probability to migrate domestically. As such, the i-MPI can also be useful for targeting purposes as it can anticipate migration decisions. It is also possible to use the scores at an aggregated level if representative data are available across broader areas.

Similar to the MPI tools calibrated for external migration, the i-MPI should be regarded as a dynamic instrument, developed for predictive purposes using the best currently available data, but whose precision can be further validated and improved as more data is collected. With the available data, the proposed model correctly identifies 81% and 84% of migrating and non-migrating households in, respectively, Guatemala and Honduras.

On a final note, it is worth clarifying that we do not intend to claim that the 10 factors included in the i-MPI are the only or even the main factors behind an individual's decision to migrate. The suit of MPI indices are just indirect measuring tools, empirically calibrated to track the probability of an individual within a household migrating in the near future. Similarly, these factors should not be interpreted as causal factors underlying migration decisions, but rather as observable markers that simply serve as proxies for identifying households with a high risk of migration. In this sense, the MPI indicators are not meant to inform decisions on development policies. Rather, comprehensive programs that tackle the multiple roots underlying migration, providing economic opportunities in both rural and urban settings, reducing crime and corruption, and overall raising the quality of life of individuals, remain the main avenue to develop resilient communities in which the incentives to migrate are reduced.

References

Ahluwalia, M.S. (2000). State-level performance under economic reforms in India. Paper presented at the Centre for Research on Economic Development and Policy Reform Conference on Indian Economic Prospects: Advancing Policy Reform. Stanford University

Almanzar, M., Ceballos, F., Constenla, S., & Hernandez, M.A. (2022). Migration Propensity Index: Calibration for Honduras. IFPRI Technical Note (internal report). Washington, DC: International Food Policy Research Institute (IFPRI)

Alverio, G.N., Sowers, J., & Weinthal, E. (2023). Peaceworks: Displaced to Cities—Conflict, Climate Change, and Rural-to-Urban Migration (No. 189).

Barrutia Barreto, I., Toledo Espinoza, E.S., Toledo Espinoza, S.L., & Seminario Córdova, R.A. (2022). Internal Migration, Ethnicity, and Acculturation Strategies in Peru 2020. In E. Alaverdov & M. Bari (Eds.), Handbook of Research on Ethnic, Racial, and Religious Conflicts and Their Impact on State and Social Security.

Barrientos, A. & Dejong, J. (2006). Reducing child poverty with cash transfers: a sure thing? Development Policy Review, 24, 537–52.

Bernard, A., & Perales, F. (2021). Linking internal and international migration in 13 European countries: Complementarity or substitution? Journal of Ethnic and Migration Studies, 1–21

Cantor, D.J., & Apollo, J.O. (2020). Internal Displacement, Internal Migration, and Refugee Flows: Connecting the Dots. Refugee Survey Quarterly, 39(4), 647–664.

Cattaneo, A., & Robinson, S. (2019). Economic Development and the Evolution of Internal Migration: Moving in Steps, Returnees, and Gender Differences. FAO Agricultural Development Economics Working Papers, No. 2019/03.

Ceballos, F., & Hernandez, M.A. (2020). The Migration Propensity Index: An application to Guatemala. IFPRI Discussion Paper 1953. Washington, DC: International Food Policy Research Institute (IFPRI).

Clemens, M. (2021). Violence, development, and migration waves: Evidence from Central American child migrant apprehensions. Journal of Urban Economics 124: 103355.

Clement, V., Rigaud, K.K., de Sherbinin, A., Jones, B., Adamo, S., Schewe, J., Sadiq, N., & Shabahat, E. (2021). Groundswell Part 2: Acting on Internal Climate Migration. World Bank.

Cohn, D., Passel, J., & Gonzalez-Barrera, A. (2017). Rise in U.S. Immigrants From El Salvador, Guatemala and Honduras Outpaces Growth from Elsewhere. Pew Research Center, December 7, https://www.pewre- search.org/hispanic/2017/12/07/rise-in-u-s-immigrants-from-el-salvador-guatemala-and-honduras-outpaces- growth-from-elsewhere/ (accessed December 2023).

Colmer, J. (2021). Temperature, Labor Reallocation, and Industrial Production: Evidence from India. American Economic Journal: Applied Economics, 13(4), 101-124.

Congressional Research Service. (2019a). Recent Migration to the United States from Central America: Frequently Asked Questions. CRS REPORT R45489, January 29, https://fas.org/sgp/crs/row/R45489.pdf (accessed December 2022).

Congressional Research Service. (2019b). Central American Migration: Root Causes and U.S. Policy. IN FOCUS IFIII5I, June 13, https://fas.org/sgp/crs/row/IF11151.pdf (accessed December 2022).

Creatives Associates International. (2019). Why Migrants Risk it All: What Drives Unauthorized Migration to the U.S.? What Shapes the Decision to Migrate? Mimeo, July.

Dallmann, I., & Millock, K. (2017). Climate Variability and Inter-State Migration in India. CESifo Economic Studies, 63(4), 560–594.

de Brauw, A., & Giles, J. (2006). Migrant Opportunity and the Educational Attainment of Youth in Rural China (IZA DP No. 2326). Discussion Paper Series. Institute for the Study of Labor.

de Brauw, A., Mueller, V., & Lee, H.L. (2014). The Role of Rural–Urban Migration in the Structural Transformation of Sub-Saharan Africa. World Development, 63(C), 33-42.

Diepart, J.C., & Ngin, C. (2020). Internal migration in Cambodia. In M. Bell, A. Bernard, E. Charles-Edwards, & Y. Zhu (Eds.), Internal migration in the countries of Asia (pp. 137–162). Springer.

Dillon, A., Mueller, V., & Salau, S. (2011). Migratory Responses to Agricultural Risk in Northern Nigeria. American Journal of Agricultural Economics, 93(4), 1048–1061.

Gray, C., & Mueller, V. (2012a). Drought and Population Mobility in Rural Ethiopia. World Development, 40(1), 134–145.

Gray, C., & Mueller, V. (2012b). Natural disasters and population mobility in Bangladesh. Proceedings of the National Academy of Sciences, 109(16), 6000-6005.

Hernandez, M.A., Ecker, O., Läderach, P., & Maystadt, J.F. (2023). Forced migration, fragility, and policy responses for developing countries. Forthcoming Chapter in 2023 Global Food Policy Report, International Food Policy Research Institute, Washington DC.

Hoogeveen, J.G., Rossi, M., & Sansone, D. (2019). Leaving, Staying or Coming Back? Migration Decisions During the Northern Mali Conflict. Journal of Development Studies, 55(10), 2089–2105.

Hosmer, D.W., & Lemeshow, S. (2000). Applied Logistic Regression (2nd Edition). John Wiley & Sons.

Ibáñez, A.M., Quigua, J., Romero, J., & Velásquez, A. (2022). Responses to Temperature Shocks: Labor Markets and Migration Decisions in El Salvador. IDB Publications (Working Papers) 12201. Inter-American Development Bank.

Institute for Economics and Peace. (2021). Ecological Threat Report 2021: Understanding Ecological Threats, Resilience and Peace. Institute for Economics and Peace.

Joulaei, H., Keshani, P., Foroozanfar, Z., Afrashteh, S., Hosseinkhani, Z., Mohsenpour, M. A., Moghimi, G., & Homayouni Meymandi, A. (2023). Food insecurity status and its contributing factors in slums' dwellers of southwest Iran, 2021: A cross-sectional study. Archives of Public Health, 81(1), 38.

McAuliffe, M., & Triandafyllidou, A. (Eds.). (2021). World Migration Report 2022. International Organization for Migration (IOM), Geneva.

National Intelligence Council's Strategic Futures Group. (2021). Structural Drivers of the Future: Demographic Trends - The Future of Migration. April 2021.

Lozano Ascencio, F. (2004). Migration Strategies in Urban Contexts: Labor Migration from Mexico City to the United States. Migraciones internacionales, 2(3), 34–59.

Lucas, R.E.B. (2016). Internal Migration in Developing Economies: An Overview of Recent Evidence. Geopolitics, History, and International Relations, 8(2), 159–191.

Ray, M., & Dutta, A. (2019). Economic reform, uneven regional development, and internal migration in India. Migration and Development, 1–20.

Raymer, J., & Baffour, B. (2018). Subsequent Migration of Immigrants Within Australia, 1981–2016. Population Research and Policy Review, 37(6), 1053–1077.

Rubenstein, J.M. (2017). Chapter 3: Migration. In The Cultural Landscape: An Introduction to Human Geography (12th ed.). Pearson.

Swinkels, R.A., Norman, T., Blankespoor, B., Munditi, N., & Zvirereh, H. (2019). Analysis of Spatial Patterns of Settlement, Internal Migration, and Welfare Inequality in Zimbabwe.

Todaro, M.P. (1969). A model of labor migration and urban unemployment in less developed countries. American Economic Review, 59 (1), 138–148.

Todaro, M.P., & Harris, J.R. (1970). Migration, unemployment, and development: A two-sector analysis. American Economic Review, 60 (1), 126–142.

United Nations High Commissioner for Refugees (UNHCR). (2019). Global Trends 2019.

Vargas-Lundius, R., Chowdhury, A., Yi, S., Seshan, G., & Yameogo, N. D. (2018). Internal Migration and Urbanization: Synthesis Note. KNOMAD.

Williamson, J.G. (1988). Migration and urbanization. In Chenery, H. and Srinivasan, T.N., (Eds). Handbook of Development Economics (pp.426–465). North-Holland: Elsevier.

Yamada, G. (2010). Patrones de migración interna en el Peru reciente. In C. Garavito, & I. Muñoz Portugal (Eds.), Empleo y Proteccion Social. Pontificia Universidad Catolica del Peru, Fondo Editorial. [In Spanish]

Appendix A

Table A.1: Universe of variables considered for the model selection in Guatemala

#	Name	Type	Definition	Timing
1	Gender of household head	Binary	One if the household head is male; zero if female	At time of the interview
2	Age of household head	Integer	Age of the household head in years	At time of the interview
3	Square of the age of household head	Integer	Square of the age of household head	At time of the interview
4	Education of household head	Integer	Number of completed years of education of the household head	At time of the interview
5	Square of education of household head	Integer	Square of the completed years of education of the household head	At time of the interview
6	Indigenous household	Binary	One if the interviewee considers him or herself to be indigenous; zero otherwise	At time of the interview
7	Land holdings	Continuous	Total hectares of agricultural land worked by a household; zero if none	During the last 12 months
8	Household has agricultural land	Binary	One if the household works any agricultural land; zero otherwise	During the last 12 months
9	Household total agricultural land is above 1000m2	Binary	One if the household works agricultural land of 1000 squared meters or above; zero otherwise	During the last 12 months
10	Produces a subsistence crop	Binary	One if the household cultivates maize or beans; zero otherwise	During the last 12 months
11	Produces a high-value crop	Binary	One if the household cultivates coffee, cardamom, a horticultural crop, or other crops apart from maize and beans; zero otherwise	During the last 12 months
12	House ownership	Binary	One if a household member is the owner of the house; zero otherwise	At time of the interview
13	Household size	Integer	Number of household members	At time of the interview
14	Number of children aged 0-5	Integer	Number of children aged 0-5 years.	At time of the interview
15	Household members aged 15-29 in household	Binary	One if the household contains at least one (male or female) member aged 15-29 years; zero otherwise	At time of the interview
16	Household members aged 30-49 in household	Binary	One if the household contains at least one (male or female) member aged 30-49 years; zero otherwise	At time of the interview
17	Household members aged 50-59 in household	Binary	One if the household contains at least one (male or female) member aged 50-59 years; zero otherwise	At time of the interview
18	Household members aged 60 or above in household	Binary	One if the household contains at least one (male or female) member aged 60 or above years; zero otherwise	At time of the interview

19	Women with no primary education	Binary	One if the household contains at least one female member with no primary education; zero otherwise	At time of the interview
20	Women with primary education or above	Binary	One if the household contains at least one female member with one year of primary education or above; zero otherwise	At time of the interview
21	Women with complete primary education or above	Binary	One if the household contains at least one female member with six years of primary education or above; zero otherwise	At time of the interview
22	Women with secondary education or above	Binary	One if the household contains at least one female member with one or more complete years of secondary education; zero otherwise	At time of the interview
23	Women members with higher education or above	Binary	One if the household contains at least one female member with one or more complete years of higher education; zero otherwise	At time of the interview
24	Men with no primary education	Binary	One if the household contains at least one male member with no primary education; zero otherwise	At time of the interview
25	Men with primary education or above	Binary	One if the household contains at least one male member with one year of primary education or above; zero otherwise	At time of the interview
26	Men with complete primary education or above	Binary	One if the household contains at least one male member with six years of primary education or above; zero otherwise	At time of the interview
27	Men with secondary education or above	Binary	One if the household contains at least one male member with one or more complete years of secondary education; zero otherwise	At time of the interview
28	Men with higher education or above	Binary	One if the household contains at least one male member with one or more complete years of higher education; zero otherwise	At time of the interview
29	Dwelling has finished walls	Binary	One if the dwelling walls are made of brick or block; zero otherwise	At time of the interview
30	Dwelling has finished ceiling	Binary	One if the dwelling ceiling is made of concrete or tiles; zero otherwise	At time of the interview
31	Dwelling has finished floors	Binary	One if the dwelling floors are made of cement, brick, ceramics, polished wood, or boards; zero otherwise	At time of the interview
32	Majority of households in the community have water supply	Binary	One if majority of households in the community obtains water for consumption from pipes in their dwelling or land; zero otherwise	At time of the interview
33	Majority of households in the community connected to electricity	Binary	One if majority of households in the community are connected to an electricity supply network; zero otherwise	At time of the interview
34	Majority of households in the community connected to sewerage	Binary	One if majority of households in the community are connected to sewerage; zero otherwise	At time of the interview
35	Household affected by excess rain or flood	Binary	One if the household has been affected by excess rain or flood, during the last 3 years; zero otherwise	During the last 3 years

36	Household affected by earthquake or volcano	Binary	One if the household has been affected by an earthquake or volcano event, during the last 3 years; zero otherwise	During the last 3 years
37	Household affected by drought	Binary	One if the household has been affected by a drought, during the last 3 years; zero otherwise	During the last 3 years
38	Household affected by frosts	Binary	One if the household has been affected by frosts, during the last 3 years; zero otherwise	During the last 3 years
39	Household owns livestock	Binary	One if the household owns any livestock; zero otherwise	At time of the interview
40	Household owns a cellphone	Binary	One if the household owns a cellphone; zero otherwise	At time of the interview
41	Household owns a television	Binary	One if the household owns a television; zero otherwise	At time of the interview
42	Household has cable television	Binary	One if the household has a cable television subscription; zero otherwise	At time of the interview
42	Household owns a radio	Binary	One if the household owns a radio; zero otherwise	At time of the interview
43	Household owns a fridge	Binary	One if the household owns a fridge; zero otherwise	At time of the interview
44	Household owns a vehicle (car/motorcycle)	Binary	One if the household owns a vehicle (car/motorcycle); zero otherwise	At time of the interview
45	Household owns houses or land	Binary	One if the household owns any real estate property; zero otherwise	At time of the interview
46	Household owns agricultural machinery or equipment	Binary	One if the household owns any agricultural machinery or equipment; zero otherwise	At time of the interview
47	Household purchased durable goods during last 12 months	Binary	One if the household made any expenditures on new durable goods during the last 12 months; zero otherwise	During the last 12 months
48	Household purchased furniture during last 12 months	Binary	One if the household made any expenditures on new furniture during the last 12 months; zero otherwise	During the last 12 months
49	Household purchased vehicles in the last 12 months	Binary	One if the household had any expenditures on vehicles in the last 12 months; zero otherwise	During the last 12 months
50	Household made home improvement during last 12 months	Binary	One if the household made any expenditures on construction, expansion, or improvement of their dwelling during the last 12 months; zero otherwise	During the last 12 months
51	Household is beneficiary of “Bolsa Segura”	Binary	One if household is beneficiary of social program “Bolsa Segura”; zero otherwise	During the last 12 months
52	Household is beneficiary of “Bono Seguro de Salud”	Binary	One if household is beneficiary of social program “Bono Seguro de Salud”; zero otherwise	During the last 12 months
53	Household has a migrant abroad	Binary	One if the household received remittances during the last 12 months; zero otherwise	During the last 12 months
54	Community-level rate of migrants abroad	Fraction [0-1]	Share of households in a household's community that received remittances during the last 12 months	During the last 12 months

55	Community-level international migration during the last 12 months	Fraction [0-1]	Share of households in the community that report a household member has emigrated abroad (to another country) since the last survey (previous 12 months)	During the last 12 months
56	Community-level internal (within country) migration during the last 12 months	Fraction [0-1]	Share of households in the community that report a household member has migrated within the country (to another department or to Guatemala City) since the last survey (previous 12 months)	During the last 12 months
57	Rural indicator	Binary	One if the household is located in a rural area according to the INE municipality-level definition; zero if urban	Static

Appendix B

Table B.1: Universe of variables considered for the model selection in Honduras

#	Name	Type	Definition	Timing
1	Gender of household head	Binary	One if the household head is male; zero if female	At time of the interview
2	Age of household head	Integer	Age of the household head in years	At time of the interview
3	Square of the age of household head	Integer	Square of the age of household head	At time of the interview
4	Education of household head	Integer	Number of completed years of education of the household head	At time of the interview
5	Square of education of household head	Integer	Square of the completed years of education of the household head	At time of the interview
6	Indigenous household	Binary	One if the interviewee considers him or herself to be indigenous; zero otherwise	At time of the interview
7	Household size	Integer	Number of household members	At time of the interview
8	Number of children aged 0-5	Integer	Number of children aged 0-5 years.	At time of the interview
9	Children aged 0-5 in household	Binary	One if the household contains at least one child aged 0-5 years; zero otherwise	At time of the interview
10	Household members aged 15-29 in household	Binary	One if the household contains at least one (male or female) member aged 15-29 years; zero otherwise	At time of the interview
11	Household members aged 30-49 in household	Binary	One if the household contains at least one (male or female) member aged 30-49 years; zero otherwise	At time of the interview
12	Household members aged 50-59 in household	Binary	One if the household contains at least one (male or female) member aged 50-59 years; zero otherwise	At time of the interview
13	Household members aged 60 or above in household	Binary	One if the household contains at least one (male or female) member aged 60 or above years; zero otherwise	At time of the interview
14	Household members with no primary education	Binary	One if the household contains at least one (male or female) member with no primary education; zero otherwise	At time of the interview
15	Household members with primary education or above	Binary	One if the household contains at least one (male or female) member with one year of primary education or above; zero otherwise	At time of the interview
16	Household members with complete primary education or above	Binary	One if the household contains at least one (male or female) member with six years of primary education or above; zero otherwise	At time of the interview

17	Household members with secondary education or above	Binary	One if the household contains at least one (male or female) member with one or more complete years of secondary education; zero otherwise	At time of the interview
18	Household members with higher education or above	Binary	One if the household contains at least one (male or female) member with one or more complete years of higher education; zero otherwise	At time of the interview
19	Dwelling has finished walls	Binary	One if the dwelling walls are made of brick or block; zero otherwise	At time of the interview
20	Dwelling has finished ceiling	Binary	One if the dwelling ceiling is made of concrete or tiles; zero otherwise	At time of the interview
21	Dwelling has finished floors	Binary	One if the dwelling floors are made of cement, brick, ceramics, polished wood, or boards; zero otherwise	At time of the interview
22	Dwelling connected to sewerage	Binary	One if the dwelling is connected to sewerage; zero otherwise	At time of the interview
23	Dwelling connected to water	Binary	One if the household obtains drinking water from public or private pipes in their own dwelling; zero otherwise	At time of the interview
24	Household has agricultural land	Binary	One if the household works any agricultural land; zero otherwise	During the last 12 months
25	Household total agricultural land is above 2mz	Binary	One if the household works agricultural land of 1000 square meters or above; zero otherwise	During the last 12 months
26	Produces a subsistence crop	Binary	One if the household cultivates maize or beans; zero otherwise	During the last 12 months
27	Produces a high-value crop	Binary	One if the household cultivates coffee, cardamom, a horticultural crop, or other crops apart from maize and beans; zero otherwise	During the last 12 months
28	Household owns livestock	Binary	One if the household owns any livestock; zero otherwise	At time of the interview
29	Household owns a cellphone	Binary	One if the household owns a cellphone; zero otherwise	At time of the interview
30	Household owns a television	Binary	One if the household owns a television; zero otherwise	At time of the interview
31	Household owns a radio	Binary	One if the household owns a radio; zero otherwise	At time of the interview
32	Household owns a fridge	Binary	One if the household owns a fridge; zero otherwise	At time of the interview
33	Household owns a vehicle (car/motorcycle)	Binary	One if the household owns a vehicle (car/motorcycle); zero otherwise	At time of the interview
34	Household purchased durable goods during last 12 months	Binary	One if the household made any expenditures on new durable goods during the last 12 months; zero otherwise	During the last 12 months
35	Household purchased furniture during last 12 months	Binary	One if the household made any expenditures on new furniture during the last 12 months; zero otherwise	During the last 12 months
36	Household purchased appliances during last 12 months	Binary	One if the household made any expenditures on new appliances during the last 12 months; zero otherwise	During the last 12 months

37	Household made home improvement during last 12 months	Binary	One if the household made any expenditures on construction, expansion, or improvement of their dwelling during the last 12 months; zero otherwise	During the last 12 months
38	Household had a savings account during the las 12 months	Binary	One if the household held a savings account during the last 12 months; zero otherwise	During the last 12 months
39	Household received a credit during the las 12 months	Binary	One if someone in the household received a credit during the last 12 months; zero otherwise	During the last 12 months
40	Household is beneficiary of Bono 10,000	Binary	One if household is beneficiary of social program Bono 10,000; zero otherwise	During the last 12 months
41	Household is beneficiary of Bono for schooling purposes	Binary	One if household is beneficiary of social program Bono for schooling purposes; zero otherwise	During the last 12 months
42	Household affected by a mild drought	Binary	One if the household has been affected by a mild drought (SPI<-1) during the past agricultural season (May to August); zero otherwise	During the last 3 years
43	Household affected by a moderate drought	Binary	One if the household has been affected by a moderate drought (SPI<1.5) during the past agricultural season (May to August); zero otherwise	During the last 3 years
44	Household has a migrant abroad	Binary	One if the household received remittances during the last 12 months; zero otherwise	During the last 12 months
45	Municipality-level rate of migrants abroad	Fraction [0-1]	Share of households in a household's municipality that received remittances during the last 12 months	During the last 12 months
46	At least one member works on dependent job	Binary	One at least one household member worked in dependent job during the last 12 months; zero otherwise	During the last 12 months
47	At least one member works on independent job	Binary	One at least one household member worked in independent job during the last 12 months: zero otherwise	During the last 12 months
48	At least one member works on independent job: sales	Binary	One at least one household member worked in independent sales job during the last 12 months; zero otherwise	During the last 12 months
49	At least one member works on independent job: agriculture	Binary	One at least one household member worked in independent agricultural job during the last 12 months; zero otherwise	During the last 12 months
50	At least one member works on independent job: manufacture	Binary	One at least one household member worked in independent manufacture job during the last 12 months; zero otherwise	During the last 12 months
51	At least one member works on independent job: professional	Binary	One at least one household member worked in independent professional job during the last 12 months; zero otherwise	During the last 12 months
52	Dependency ratio	Ratio	Number of dependent household members (under 14 and over 65 years of age) divided by the number of household members of working age (from 15 to 64 years of age)	At the time of the interview
53	Active working members to dependents ratio	Ratio	Number of household members that worked in the last 12 months divided by the number of	During the last 12 months

			dependent household members (under 14 and over 65 years of age)	
54	Active working members to members of working age ratio	Ratio	Number of household members that worked in the last 12 months divided by the number of household members of working age (from 15 to 64 years of age)	During the last 12 months
55	Remoteness	Numeric	Distance from the village where the household is located to the nearest big market or population area	Static
56	House ownership	Binary	One if a household member is the owner of the house; zero otherwise	At time of the interview
57	Land property	Binary	One if a household member owns agricultural land; zero otherwise	At time of the interview
58	Community-level rate of households with finished walls	Fraction [0-1]	Share of households in the community that have finished walls; zero otherwise	At time of the interview
59	Community-level rate of households with finished ceiling	Fraction [0-1]	Share of households in the community that have finished ceiling; zero otherwise	At time of the interview
60	Community-level rate of households with finished floors	Fraction [0-1]	Share of households in the community that have finished floors; zero otherwise	At time of the interview
61	Community-level rate of households connected to electricity supply	Fraction [0-1]	Share of households in the community that is connected to an electricity supply network; zero otherwise	At time of the interview
62	Community-level internal (within country) migration during the last 12 months	Fraction [0-1]	Share of households in the community that report a household member has migrated within the country (to another department or to Tegucigalpa) since the last survey (previous 12 to 24 months)	During the last 12 to 24 months
63	Household member with a permanent disability	Binary	One if a household member has a permanent disability; zero otherwise	At time of the interview
64	Rural indicator	Binary	One if the household is located in a rural area; zero if urban	Static

Note: SPI = Standard Precipitation Index. A household is considered to have suffered a drought event if for any of the months that comprise the last *primera* agricultural season prior to the moment of the survey, the SPI was lower than -1 (in the case of the mild drought variable) or lower than -1.5 (in the case of the moderate drought variable). The monthly SPI was estimated using 20 years of historic precipitation data (2000-2020), at the municipality level, from the Climate Hazards Group InfraRed *Precipitation* with Station data (*CHIRPS*).

Appendix C

Table C.1: i-MPI score-to-probability look up table

i-MPI score	Probability	Risk category	i-MPI score	Probability	Risk category	i-MPI score	Probability	Risk category
24 or under	< 0.001	LOW RISK	50	0.010	MEDIUM RISK	76	0.186	HIGH RISK
25	0.001		51	0.011		77	0.205	
26	0.001		52	0.013		78	0.225	
27	0.001		53	0.014		79	0.246	
28	0.001		54	0.016		80	0.269	
29	0.001		55	0.018		81	0.294	
30	0.001		56	0.020		82	0.319	
31	0.001		57	0.023		83	0.346	
32	0.001		58	0.026		84	0.373	
33	0.001		59	0.029		85	0.402	
34	0.001		60	0.032		86	0.431	
35	0.002		61	0.036		87	0.460	
36	0.002		62	0.041		88	0.490	
37	0.002		63	0.046		89	0.520	
38	0.002		64	0.051		90	0.550	
39	0.003		65	0.058		91	0.579	
40	0.003		66	0.064		92	0.608	
41	0.003		67	0.072		93	0.637	
42	0.004		68	0.080		94	0.664	
43	0.004		69	0.090		95	0.690	
44	0.005		70	0.100	HIGH RISK	96	0.715	
45	0.006		71	0.111		97	0.739	
46	0.006		72	0.124		98	0.761	
47	0.007		73	0.137		99	0.782	
48	0.008		74	0.152		100 or above	> 0.800	
49	0.009		75	0.168				