Spillover effects in adoption of Cash Transfer Programs by Latin American countries

Diego E. Vacaflores
Texas State University
Department of Finance & Economics
601 University Drive
San Marcos, TX 78666, USA
E-mail: dv13@txstate.edu
and
James P. LeSage
Fields Endowed Chair
Texas State University
Department of Finance & Economics
601 University Drive
San Marcos, TX 78666, USA
E-mail: james.lesage@txstate.edu

August 2, 2019

Abstract

Some of the most effective public programs used in Latin America to reduce poverty and inequality have been non-contributory cash transfers. We examine country-specific characteristics that lead countries to adopt these programs over time using a state-transition spatial probit panel data model that takes into account dependence between countries' decision to adopt these programs. Intuitively, past adoption of cash transfer programs by other countries might have an impact on the probability that a country implements this type of program. We explore alternative connectivity structures to model dependence, spatial proximity as well as connections based on population migration flows, finding out-migration as most consistent with our sample data and spatial regression specification. For our panel of 17 Latin American countries over the period 2000 to 2017 we find evidence of dependence between countries in the probability of adoption of conditional cash transfer programs (CCT), but no such evidence in the case of unconditional cash transfer programs (UCT).

KEYWORDS: Conditional Cash Transfer programs, Government Programs, Latin America, Cross-sectional Dependence.

JEL: I38; N96; C51; R11.

1 Introduction

Latin American countries have been using non-contributive cash transfer programs as redistributive schemes to alleviate poverty and promote economic activity since the late 1990s. These programs mimick the *Prospera* initiative in Mexico and *Bolsa Familia* in Brazil. Some of these programs are conditioned on participation in human capital development efforts, like schooling and health care checkups, which we label *conditional cash transfer* programs (CCT), while others have no strings attached, simply transferring resources between segments of the population, which we label *unconditional cash transfer* programs (UCT). By 2013, CCT programs reached 135 million people in 17 Latin American and Caribbean countries while UCT schemes benefited 17 million individuals in these countries. Beneficiaries accounted for approximately 90 percent of the poor in the case of CCT transfers – although these programs reached only half of the extremely poor.[1]

Figure 1: Cash transfer programs as a percent of GDP in Latin America

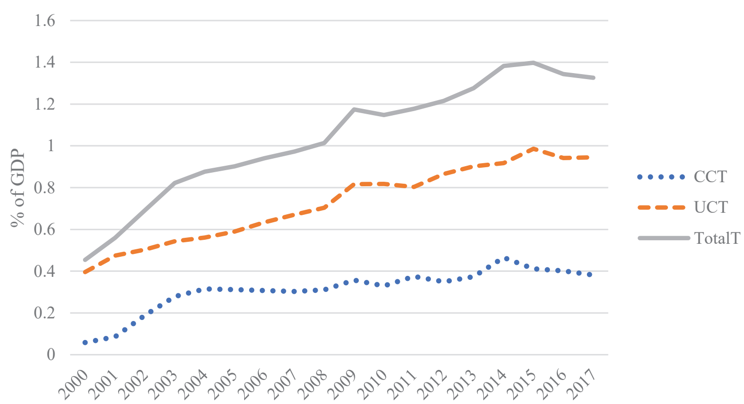
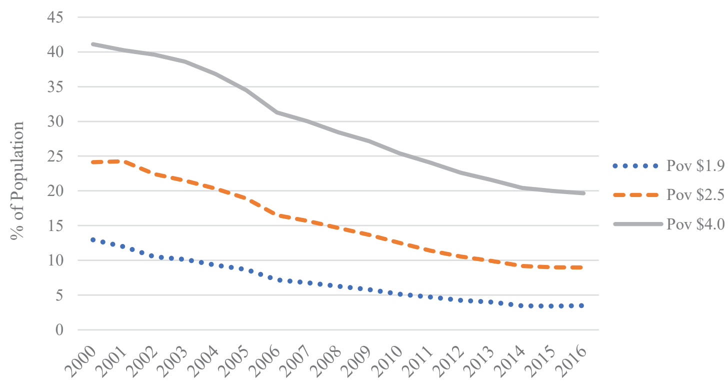


Figure ?? shows the percent of GDP devoted to cash transfer programs in our sample of 17 Latin American countries during the 2000-2017 time period. While only Brazil, Ecuador and Mexico had CCT programs in place in 2000, devoting an average of 0.29 percent of GDP in each country (which represented 0.05 percent of total regional GDP), the popularity of this type of programs led thirteen other countries to offer CCT transfers by 2017 – devoting approximately $20.4 billion US dollars, or almost 0.4 percent of regional GDP. Only four countries provided UCT transfers in 2000, accounting for almost 0.4 percent

[1] The distinction between poor and extremely poor is that those earning less than $2.5 dollars a day are considered extremely poor while those earning less than $4 dollars a day are poor (experiencing moderate poverty).

of regional GDP, but the growth of these programs has increased funding to $50.6 billion US dollars (almost 0.95 percent of regional GDP), when fourteen of the countries in our sample had these programs in place. Overall public transfers (UCT plus CCT) accounted for approximately 1.3 percent of regional GDP in 2017.

Figure 2: Poverty rates in Latin America



While there are no long-term studies of the influence these cash transfers can have on poverty and productivity, evaluation of specific programs in some countries have provided substantial evidence of a beneficial impact, leading policymakers and scholars to promote these programs because they seem to effectively reduce poverty and promote equality. Figure **??** shows that poverty in the 17 Latin American countries of our sample has declined from approximately 41 percent in 2000 to 20 percent in 2017 (lifting approximately 112 million people out of poverty). These programs aim to raise labor market skills and health of beneficiaries, leading to enhanced productivity, better paying and more secure jobs that move families out of poverty. The early success of such programs has led to the implementation of similar programs in more than sixty countries around the world.

The literature provides evidence of program success – especially for the CCT programs which have been evaluated – in term of improvements in educational attainment, nutrition, consumption, and labor market participation. Parker and Todd (2017) for example note that the *Oportunidades* program (now renamed *Prospera*) was able to reduce poverty while improving school attendance, grade progression, health care access, savings and even income. Schultz (2004) and Berhman et al. (2005) for their part find that the Mexican program has improved school enrollment and facilitated grade progression. Furthermore, Todd and Wolpin (2006) show that conditional cash transfers are significantly more efficient in improving schooling, relative to unconditional transfers. But cash transfer programs do not only affect education and health care access, they can also impact economic engagement of participating households. For example, Bianchi and Bobba (2013) examine the behavior of entrepreneurs in Mexico and find that recipients of these transfers increase risk-taking because they provide a stable source of income. Recipients of cash transfers showed a greater willingness to start self-employed ventures, increasing micro-entrepreneurship. In addition, Gertler et al. (2012) show that recipients of these monetary transfers were able to increase their long-term income and consequently raise their consumption levels (presumably by investing part of these transfers in productive initiatives). Indeed, Berham et al. (2011) and Parker and Vogl (2018) find that *Oportunidades* raised education and labor force participation of females in the longer-term.

Aizer et al. (2016) in an empirical study show that unconditional cash transfers in the U.S. between 1911 and 1935 produced long term improvements in longevity, education, health, and income. Other studies have used theoretical models to overcome data limitations and corroborate long-run beneficial impacts. For example, Cespedes (2014) found an increase in human capital and years of education using a simulation model for the Mexican program, reducing poverty and income inequality in the long-run and fueling an economic expansion of approximately 6.5 percent. Peruffo and Calvanti (2017) for their part calibrate their model to the Brazilian *Bolsa Familia* program, and find that conditional cash transfers have a significant effect on increasing primary school educational attainment and reducing child labor in the long-run, although it temporarily forces children to work more to became eligible to participate in the program. Human capital slowly builds over time leading to future increases in output. Vacaflores (2019) provides a model that allows increases in cash transfers to lower poverty rates and create economic growth if the program raises productivity by a large enough margin. His results are based on a model calibrated to the same 17 Latin American countries used in this study.

While these studies provide insight into the effectiveness of these cash transfers on schooling, nutrition, health care and productivity, very little is known about factors that lead to implementation of such programs. We hypothesize that countries would be more inclined to implement redistributive cash transfer programs when they experience high levels of poverty, or inequality, and when they have enough resources to support these redistribute schemes. Another hypothesis is that in the presence of cross-sectional dependence, effectiveness of such programs in one country should exert spillover impacts on the probability of implementation of such programs in other countries.

We contribute to the literature on cash transfers by exploring country-level characteristics associated with the probability of adoption of UCT and CCT programs. Our approach relies on a state-transition panel probit model that allows for cross-sectional dependence in decisions to adopt UCT and CCT programs by individual countries. This involves estimation of the model using state-transition behavior of 17 Latin American countries some of which have implemented these programs during the 2000 to 2017 period. By state-transition we mean that countries adopting the program transition from a 0-state to a 1-state at time t when they adopt the program.[2] The sequence of 0,1 states for our panel of 17 countries over time represents the dependent variable that we model using a panel probit specification that allows for cross-country dependence. The specification allows the 0,1 state of one country to depend on program implementation decisions made by other countries.

We explore alternative *exogenous* specifications for the connectivity structure that describes the dependence relationships. One specification defines the dependence of country i's decision to be based on decisions made by the set $j \in \mathcal{S}$ of geographically neighboring countries (those with common borders). Another defines decision dependence of country i as based on the set $j \in \mathcal{I}$ of countries that provide the largest number of population in-migrants, and a final specification defines country i's decision as dependent on the set $j \in \mathcal{O}$ of countries to which country i sends a large number of population out-migrants.

We use data on Conditional Cash Transfers and disbursements to Non-Contributory Pension schemes gathered by the Economic Commission for Latin America and the Caribbean (ECLAC) to construct a cross-sectional panel of countries that covers the 2000-2017 time period, and we note these transfers do not include social security payments. We explore how country-specific characteristics impact the probability of CCT and UCT program adoption, and test for the presence of significant spillover effects arising from dependence of program adoption decisions made by each country on a set of j other countries. Intuitively, factors such as the size and role of the government sector in the economy, the extent of population living in poverty, population size, and levels of economic development would have an impact on the decision to adopt cash transfer programs. By spillover effects we mean that changes

[2] Elhorst et al. (2017) propose this methodology and apply it to adoption of inflation targeting regimes by a sample of countries.

in country i characteristics will have an impact on the probability that country i adopts the program as well as an impact on the probability that countries in the set j also adopt the program. Our model quantifies the magnitude of these spillover impacts resulting from dependence in program adoption decisions.

We find evidence of positive dependence of program adoption decisions for our sample of Latin American countries in the case of CCT programs, but no significant dependence in adoption decisions regarding the UCT programs. Positive dependence in the probability of CCT program adoption suggests that the presence of CCT programs in the set j of peer countries increases the probability of program adoption in the typical country i.[3] The lack of significant dependence of (the typical) country i decisions regarding adoption of UCT programs suggests that the presence or absence of UCT programs in the set j of other countries has no influence on the probability of adopting a UCT program.

Section **??** develops the panel probit state-transition model specification that allows for decision-dependence regarding program adoption between countries. We provide a theoretical motivation for this model, as well as a motivation for three alternative dependence structures that exogenously specify the dependence sets $j \in \mathcal{S}, j \in \mathcal{I}$ and $j \in \mathcal{O}$. We also discuss log-probability and quadratic probability scores (LPS, QPS) that are used to determine which of the alternative dependence structures is most consistent with the model specification and sample data. Section **??** contains a description of Markov Chain Monte Carlo (MCMC) estimation of the state-transition panel probit model along with a discussion of how to interpret estimates and draw inferences for this model. Section **??** presents results from application of the model to the panel of 17 Latin American countries. Section **??** concludes.

2 A state-transition panel probit model for program adoption

LeSage and Pace (2009, Chapter 10) set forth a spatial Durbin model (SDM) variant of the conventional cross-sectional probit model that we can adapt to the (static) panel data case here. In (**??**), $y^\star$ represents an $NT \times 1$ vector reflecting the latent unobservable utility

[3] Like all regression models, inferences from our model should be interpreted as reflecting an average across the sample of countries.

associated with each country i adopting a transfer program at time t, and $\otimes$ represents the Kronecker product. We let the matrix $W(j)$ denote an $N \times N$ matrix with non-zero values in the i, jth position for countries in the set j, where j varies according to the set of countries on which there is decision dependence ($j \in \mathcal{S}, j \in \mathcal{I}, j \in \mathcal{O}$), respectively. The matrices $W(j)$ are row-normalized to have row-sums of unity, and $\varepsilon \sim N(0, I_{NT})$.

$$y^{\star} = \rho(I_T \otimes W(j))y^{\star} + X\beta + (I_T \otimes W(j))X\theta + \iota_T \otimes \mu_N + \iota_N \otimes \mu_T + \varepsilon. \qquad (1)$$

We use an $N \times T$ matrix Y with zero, one values to reflect the (observable) presence or absence of the transfer program in country $i = 1, \ldots, N$ at time $t = 1, \ldots, T$. The observable dependent variable in our model is then $y = vec(Y)$, an $NT \times 1$ vector containing zero, one values. The $NT \times k$ matrix X contains explanatory variables consisting of time-varying country-specific characteristics for each country at each time period. The scalar parameter ρ measures the strength of dependence, with a value of zero indicating independence. Clearly, a conventional panel probit model reflecting decisions that are independent emerges when $\rho = 0$. It should be noted that we have a *static* panel data model which means that the *same* dependence structure exists between countries for all time periods. The block diagonal matrix $(I_T \otimes W(j))$ does not allow for interaction between different time periods, only simultaneous dependence at each time period t. The estimate for the scalar parameter ρ reflects an average level of dependence over all countries and time periods, and the static model implies the same data generating process (DGP) operates for all time periods.

The model relationship also includes characteristics of the set j of countries on which decisions are dependent, denoted by the Kronecker matrix product involving the explanatory variables, $(I_T \otimes W(j))X$. These variables would reflect an average of characteristics for each country i's set j of countries defined by the dependence structure j. This set of explanatory variables can be thought of a reflecting the *context* in which policy-makers in each country operate. For example, country i may be a small/large population country with low/high income levels with a dependence set j of countries that are on average large/small population countries with high/low income levels. The model relationship allows for these contextual effects to explain variation in the utility of transfer program adoption.

The model also allows for an $N \times 1$ vector of country-specific fixed effects (μ_N) and a set of T time-specific effects μ_T.

2.1 A theoretical motivation for the state-transition panel probit model

The Bayesian approach to modeling binary limited dependent variables treats the binary 0,1 observations in y as indicators of latent, unobserved $y^\star$ (net) utility associated with the choice of adopting the transfer program, with the unobservable utility underlying the observed (0,1) pattern of program adoptions over time. For example, in our case where the binary dependent variable reflects the presence/absence of the transfer program, the decision to adopt the program would be made when the net (perceived) utility from having the transfer program ($y = 1$) versus not having the program ($y = 0$) is positive. LeSage and Pace (2009, Chapter 2) provide a number of theoretical econometric motivations for how situations arise where policy makers in one country would exhibit utility that depends on that of policy-makers in other countries.

One of those scenarios is relevant here. If program adoption decisions by politicians are posited to be influenced by behavior of politicians located in other countries in the dependence set j countries in previous time periods, then we can formally express this type of utility dependence of $y_t^\star$ at time t on past period utility of politicians in the set of countries j as shown in (**??**). Note that politicians in country i at time t can observe the presence/absence of the transfer program in the dependence set of countries j during the previous period.

$$
\begin{aligned}
y_t^\star &= G(j)y_{t-1}^\star + Z\delta + \varepsilon_t, &(2)\\
G(j) &= (\rho(I_T \otimes W(j)),\\
Z &= \begin{pmatrix} X & (I_T \otimes W(j)) \end{pmatrix},\\
\delta &= \begin{pmatrix} \beta & \theta \end{pmatrix}'.
\end{aligned}
$$

Where we have assumed that underlying characteristics of the countries X remain relatively fixed over time or exhibit growth at a constant rate: $X_t = \phi^t X_0$, allowing us to write Z without a time subscript. Since variation in country-level characteristics such as poverty, population size, level of development, and the relevance of the public sector in the economy – measures used in the matrices X_t of our model – change slowly over time, this assumption seems reasonable. The assumption implies that country population or GDP growth (say, GDP per capita) is constant over time, and dependent on initial period size. Of course, this assumption need only be approximately valid to justify the results that follow.

The dynamic relationship in (**??**) implies a relationship for time $t-1$ shown in (**??**), which can be used to replace $y^\star_{t-1}$ in (**??**), resulting in the expressions in (**??**) and (**??**).

$$
\begin{aligned}
y^\star_{t-1} &= G(j)y^\star_{t-2} + Z\delta + \varepsilon_{t-1} && (3) \\
y^\star_t &= Z\delta + G(j)(Z\delta + G(j)y^\star_{t-2} + \varepsilon_{t-1}) + \varepsilon_t && (4) \\
y^\star_t &= Z\delta + G(j)Z\delta + G(j)^2 y^\star_{t-2} + G(j)\varepsilon_{t-1} + \varepsilon_t && (5)
\end{aligned}
$$

Recursive substitution of past values for the vector $y^\star_{t-r}$ on the right-hand size of (**??**) over q time periods leads to (**??**).

$$
\begin{aligned}
y^\star_t &= (I_{n^2} + G(j) + G(j)^2 + \ldots + G(j)^{q-1})Z\delta + G(j)^q y^\star_{t-q} + u && (6) \\
u &= \varepsilon_t + G(j)\varepsilon_{t-1} + G(j)^2\varepsilon_{t-2} + \ldots + G(j)^{q-1}\varepsilon_{t-(q-1)}
\end{aligned}
$$

Expression (**??**) can be simplified by noting that $E(\varepsilon_{t-r}) = 0, r = 0, \ldots, q-1$, which implies that $E(u) = 0$. In addition, the magnitude of $G(j)^q y^\star_{t-q}$ becomes small for large q, given the stability restrictions for the dependence parameter $(-1 < \rho < 1)$ and the fact that the matrix $W(j)$ is row-stochastic (has row-sums of unity), since row-stochastic matrices have a principle eigenvalue of one.

The implication of this development is that we can interpret the dependence that arises in the model for time t utility as the outcome or expectation of a long-run equilibrium or steady state relationship, shown in (**??**).

$$
\begin{aligned}
\lim_{q\to\infty} E(y^\star_t) &= (I_{NT} - G(j))^{-1} Z\delta && (7) \\
&= (I_{NT} - \rho(I_T \otimes W(j)))^{-1} Z\delta
\end{aligned}
$$

This is the expectation for the time t data generating process of the static panel data probit model given in (**??**), where as noted the static panel model assumes a single DGP for all time periods.

2.2 A motivation for alternative dependence sets $W(j)$

Estimates from the model are conditional on the specific type of dependence set used to define the matrix $W(j)$ that determines the set of countries on which program adoption decisions are dependent. We consider three alternative dependence sets $j \in \mathcal{S}, j \in \mathcal{I}, j \in \mathcal{O}$, each of which defines an alternative group of countries on which program adoption decisions depend. Intuitively, this dependence is related to the way in which information propagates between countries. One hypothesis is that geographical proximity facilitates transfer of information, with information transfer decaying with distance between countries. We define the set of countries $j \in \mathcal{S}$ as those with common borders to each country i, reflecting the notion that information about program success in neighboring countries is more readily available to decision makers in country i. Another definition of the set $j \in \mathcal{I}$ was based on countries that provide a large proportion of in-migrants to each country i who bring information with them about the presence/absence and success/failure of cash transfer programs in their countries of origin. For example, in-migrants from countries with successful cash transfer programs such as Brazil or Mexico may have experienced the benefits of *Bolsa Familia* and *Prospera* and would share information regarding their value with residents of the destination country. The third definition of the set $j \in \mathcal{O}$ was based on countries to which a large proportion of out-migrants from each country i flow. Here, it could be the case that migrants transfer information about the presence/absence of social programs in their destination country back to families in their country of origin – migrants from Nicaragua living in Mexico could be sharing information about the *Prospera* program in Mexico with their families back in Nicaragua.

To identify countries in the sets $j \in \mathcal{I}, j \in \mathcal{O}$ we rely on data from *Trends in International Migrant Stock*, from the United Nations, and use the 2010 figures to determine the relative importance of each country according to their migration patterns (migration flows have been relatively stable since the turn of the century for the countries of our sample).[4] The $N \times N$ weight matrices $W(j \in \mathcal{I})$ and $W(j \in \mathcal{O})$ were constructed based on the migrant stock from each other country in the sample with zeros on the main diagonal, reflecting migration from/towards all other countries. We assign zero value to row-elements that represent less

[4] Venezuela is going through an intense migration process, but is not included in our sample.

than 8 percent of the total in- or out-migration stock to produce a number of non-zero elements in the matrices $W(j \in \mathcal{I})$ and $W(j \in \mathcal{O})$ similar to the matrix $W(j \in \mathcal{S})$ based on spatial proximity (countries with common borders).[5] Specific weight matrix elements were then defined for each country of origin or destination based on the migration stock from these countries, and the weight matrices were re-normalized to have row-sums of unity.

Because migration exhibits a 'gravity effect' migrants are deterred from moving to countries further away, so in many cases the source country of in-migrants would be immediately neighboring countries. This might also be true of destination countries for out-migrants. If this is the case, the three alternative weight matrices $W(j \in \mathcal{S}), W(j \in \mathcal{I}), W(j \in \mathcal{O})$ would exhibit a high degree of similarity. To measure similarity in the alternative dependence matrices we use an approach set forth in LeSage and Pace (2014), who suggest using the correlation between the vectors: $(I_T \otimes W(j \in \mathcal{S})y, (I_T \otimes W(j \in \mathcal{I})y, (I_T \otimes W(j \in \mathcal{O})y$, where y is our $NT \times 1$ dependent variable vector containing 0,1 values. These correlations range from a low of 0.70 up to a high of 0.87 for both programs, so we should see relatively similar estimates and inferences from model specifications based on alternative definitions of the dependence sets.

To distinguish between performance of models based on these alternative dependence set definitions we calculated log-probability and quadratic probability scores (LPS, QPS). These scores represent an analogue to mean absolute error and root mean squared error for situations where the observed outcomes are binary 0,1 and the model predictions are probabilities of the 0,1 outcomes. If we let $\hat{p}_{y_i=1}$ denote probability predictions from our model and y the observed 0,1 values, $QPS = (\sum_{i=1}^{NT} 2(\hat{p}_{y_i=1} - y_i)^2)/NT$, and $LPS = -(1/NT)\sum_{i=1}^{NT}(1-y_i)log(1-\hat{p}_{y_i=1}) + ylog(\hat{p}_{y_i=1})$. The QPS ranges from 0 to 2, with QPS scores of 0 reflecting perfect accuracy, and LPS values closer to 0 reflect better accuracy.

Table **??** shows the LPS and QPS results for models involving both UCT and CCT program adoption decisions based on the three alternative definitions of j countries that represent the dependence sets. From the results we see that the definition of the set $j \in \mathcal{O}$ is most consistent with our model specification and sample data. The relatively small differences between scores reported in the table is consistent with the high correlation noted

[5] There were 49 non-zero elements in the matrix $W(j \in \mathcal{S})$ and 55 non-zero elements in the matrix $W(j \in \mathcal{I})$ defined using the stock of in-migration flows with the 8 percent cut-off. Without the cut-off, there were 148 non-zero elements in the matrix $W(j \in \mathcal{I})$. When connectedness is defined using out-migration we ended up with 45 non-zero elements in the matrix $W(j \in \mathcal{O})$ with the 8 percent cut-off, down from the original 145 non-zero elements without the threshold.

for spatial lag vectors based on alternative $W(j)$ matrices. Based on these results, we will report results in our empirical application based on the dependence set of countries defined by out-migration ties $W(j \in \mathcal{O})$ between our sample of 17 countries. This approach greatly simplifies presentation of results, and allows us to use the matrix W to represent the weight matrix defined on the basis of $j \in \mathcal{O}$ in the sequel to simplify notation.[6]

Table 1: LPS and QPS results for alternative dependence set definitions

	CCT			UCT	
Weights	LPS	QPS		LPS	QPS
$W(j \in \mathcal{S})$	0.1079	0.0411		0.1479	0.0575
$W(j \in \mathcal{I})$	0.1108	0.0425		0.1261	0.0481
$W(j \in \mathcal{O})$	0.0911	0.0338		0.1222	0.0463

3 Estimation and interpretation of the model

The Bayesian estimation approach to these models is to replace the unobserved latent utility with *parameters* that are estimated. For the case of our SDM probit model, given estimates of the $NT \times 1$ vector of missing or unobserved (parameter) values that we denote as $y^{\star}$, we can proceed to estimate the remaining model parameters β, ρ, θ by sampling from the same conditional distributions that are used in the continuous dependent variable Bayesian SDM models (see Chapters 5 and 10 in LeSage and Pace (2009)).

There is however the issue of fixed effects for both countries and time periods that arise in our panel probit model specification. We transform the explanatory variables to deviation from means form to conform with the probit assumption that $\sigma_{\varepsilon}^2 = 1$, which should eliminate country-specific fixed effects. Time-specific effects (dummies) were included in the model as additional explanatory variables, which we subsume in the matrix $X_0 = \begin{pmatrix} X & \iota_N \otimes \mu_T \end{pmatrix}$. The matrix Z containing own-region explanatory variables, time dummies and dependence region explanatory variables consist of $Z = \begin{pmatrix} X_0 & (I_T \otimes W)X \end{pmatrix}$.[7]

More formally, the program adoption choice (at time t) depends on the difference in utility: $(\pi_{1i} - \pi_{0i}), i = 1, \dots, N$ associated with observed 0,1 program absence/presence indicators, where π_{1i} represents utility (of country i) associated with program adoption and π_{0i} that from not having the program. The probit model assumes this difference at each

[6]The full results for the three different dependence set definitions are available on the author's webpage.
[7]Of course, we do not want to transform the dependent variable vector that consists of 0,1 values.

time period t, $y^*_{it} = \pi_{1it} - \pi_{0it}$, follows a normal distribution. We do not observe y^*_{it}, only the program adoption choices made, which are reflected in:

$$
\begin{aligned}
y_{it} &= 1, \quad \text{if} \quad y^*_{it} \geq 0 \\
y_{it} &= 0, \quad \text{if} \quad y^*_{it} < 0
\end{aligned}
$$

If the vector of latent utilities y^* were known, we would also know y, which led Albert and Chib (1993) to conclude: $p(\beta, \rho, \theta | y^*) = p(\beta, \rho, \theta | y^*, y)$.[8] The insight here is that if we view y^* as an additional set of parameters to be estimated, then the (joint) conditional posterior distribution for the model parameters β, ρ, θ (conditioning on both y^*, y) takes the same form as a Bayesian regression problem involving a continuous dependent variable rather than the problem involving the discrete-valued vector y. This approach was used by LeSage and Pace (2009, Chapter 10) to implement a Bayesian Markov Chain Monte Carlo estimation procedure for the SDM probit model. We rely on this approach to estimate the parameters of our static panel data probit model.

Interpreting the way in which changes in the explanatory variables in the matrix X impact the probability of a country choosing to adopt the cash transfer program in the SDM probit models requires some care.[9] The expressions in (**??**) make it clear that the probability (of a 0,1 program adoption outcome) is a non-linear function $F(\cdot)$ (the multivariate probability rule) of a function $(I_{NT} + \rho(I_T \otimes W) + \rho^2(I_T \otimes W^2) + \ldots)Z\delta$ of own-country characteristics reflected by X, as well as country characteristics from those in the dependence set of country i captured by $(I_T \otimes W)X$. We note that since we have a static panel data relationship, the DGP and partial derivatives showing how changes in own- and connected-country characteristics impact the probability of transfer program adoption are the same for all time periods.

$$
\begin{aligned}
y &= \rho(I_T \otimes W)y + Z\delta + \varepsilon \qquad (8) \\
y &= S(\rho)Z\delta + S(\rho)\varepsilon
\end{aligned}
$$

[8] Of course, Albert and Chib (1993) did not deal with the case of spatial dependence, so $\rho = 0$ in their independent probit model.

[9] We do not interpret coefficients associated the time dummy variables in the matrix X_0, just those associated with explanatory variables in the matrix X.

$$
\begin{aligned}
Z\delta &= \sum_{v=1}^{k} x_v\beta_v + \sum_{t=1}^{T-1}(\iota_N \otimes \mu_T)\beta_t + \sum_{v=1}^{k}(I_T \otimes W)x_v\theta_v \\
S(\rho) &= (I_{NT} - \rho(I_T \otimes W))^{-1} = I_{NT} + \rho(I_T \otimes W) + \rho^2(I_T \otimes W^2) + \ldots \\
\Pr(y = 1|Z) &= F\{S(\rho)(Z\delta)\}
\end{aligned}
$$

We point the reader's attention to the definition of $Z\delta = \sum_{v=1}^{k} x_v\beta_v + \sum_{t=1}^{T-1}(\iota_N \otimes \mu_T)\beta_t + \sum_{v=1}^{k}(I_T \otimes W)x_v\theta_v$ in (**??**). This definition makes it clear that changes in the vth explanatory variable in the matrix X will result in a partial derivative involving both parameters β_v, θ_v. Since the interpretation of the cross-sectional dependence probit model builds upon the interpretation of changes of independent variables on dependent observations as well as the conventional non-linear transformations due to the probit model, to simplify the exposition we first consider the simpler case of a non-probit spatial regression model shown in (**??**), where u denotes a continuous $NT \times 1$ dependent variable vector.

$$
\begin{aligned}
u &= \alpha\iota_n + \rho W u + Z\delta + \varepsilon \qquad (9) \\
\partial u/\partial x_v' &= (I_{NT} - \rho(I_T \otimes W))^{-1}(I_{NT}\beta_v + (I_T \otimes W)\theta_v), \quad v = 1, \ldots, k \\
&= S(\rho)(I_{NT}\beta_v + (I_T \otimes W)\theta_v)
\end{aligned}
$$

We could extend the approach of LeSage and Pace (2009) to our panel probit model and use an average of the diagonal elements from the $NT \times NT$ matrix: $\partial u/\partial x_v'$ to produce a scalar summary of the *direct effects*, which are derived from the own partial derivatives: $\partial u_i/\partial x_{v,i}$. Similarly, we could use an average of the (cumulated) off-diagonal elements from the $NT \times NT$ matrix: $\partial u/\partial x_v'$ $(i \neq j)$ to produce a scalar summary of the (cumulative) *indirect effects* associated with the cross-partial derivatives: $\partial u_i/\partial x_{v,j}$. This scalar summary measure cumulates the spillovers falling on counties in the dependence set of country i as well as countries in the dependence sets of those countries in the dependence set of the countries in the set j, and so on.

When we allow for dependence among observations/countries, changes in the explanatory variables associated with one country, say poverty in county i, will influence the dependent variable value reflecting program adoption in county i as well as other counties in the set j. For the case of decision dependence, the (non-zero) cross-partial derivatives represent what are commonly thought of as spillover impacts. Changes in the value of an explanatory variable in a single observation/country i can (potentially) influence all $N-1$ other observations/countries. This is true for all $i = 1, \ldots, N$ values of the vth explanatory variable leading to an $N \times N$ matrix of own- and cross-partial derivatives. We note that given the block diagonal nature of the matrix $I_T \otimes W$, impacts are not transmitted to countries in other time periods, resulting in an $N \times N$ matrix of own- and cross-partial derivatives, that depends on estimates for the parameters ρ, β, θ which reflect averages over all countries and time periods. LeSage and Pace (2009) argue for the use of scalar summary measures of the $N \times N$ matrix of own- and cross-partial derivatives based on an average across the sample of observations, similar in spirit to the way conventional least-squares regression estimates are interpreted. Specifically, an average of the main diagonal elements from the $N \times N$ matrix reflecting own-partials is used as a scalar summary for the *direct* effects, and an average of the cumulative sums of off-diagonal elements from each row is used as a scalar summary for the *indirect or spillover* effects. An important point is that the scalar summary measure of *spillover effects* cumulates the spillovers falling on all other observations, but the magnitude of impact will be greatest for countries in the immediate dependence set j and decline in magnitude for higher order dependence.[10] The sum of the two effects (*direct* and *indirect*) represent the (cumulative) *total effect* associated with a change in an observation for that explanatory variable.

The decision dependence model collapses to an independence model when the scalar dependence parameter ρ takes a value of zero. In this case, the cross-partial derivatives reflecting spillovers are all zero. An implication of this is that conventional probit models assume independence between decisions of observations which restricts spillovers to be zero.

Turning to the more complicated case of the dependence panel probit model, consider the impact on observation/country i arising from a change in a variable x_v (say the rate of poverty) in a country in the set j, a single cross-derivative is shown in (**??**) (See LeSage et al. 2011).[11]

[10] If the set j represents countries that depend on country i, then second-order dependence would be on countries in the dependence sets of the countries in j, say the sets $k_1, k_2, \ldots k_m$ for the m countries in the set j. Third-order dependence would be on countries in the dependence sets of the countries in $k_1, k_2, \ldots k_m$, that are in the dependence sets $k_1, k_2, \ldots k_m$ of j, which is the dependence set of country i, and so on.

[11] We have extended their development to the case of a static panel data setting here.

$$
\begin{aligned}
\eta &= S(\rho)Z\delta = E(y^*) && (10)\\
\frac{\partial \Pr(y_i = 1)}{\partial x_{v,j}} &= \left(\frac{\partial F(\eta)}{\partial \eta}\Big|\eta_i\right) S_{ij}(\rho)(\beta_v + (I_T \otimes W)\theta_v)\\
\frac{\partial \Pr(y_i = 1)}{\partial x_{v,j}} &= \mathrm{pdf}(\eta_i) S_{ij}(\rho)(\beta_v + (I_T \otimes W)\theta_v) && (11)
\end{aligned}
$$

We note that if $\rho = 0$ so that $S(\rho) = I_{NT}$ (and $S_{ij}(\rho) = 0$), we arrive at the standard probit result where changes in x-values of a connected country j have no impact on country i's decision to adopt the transfer program. LeSage et al. (2011) construct a matrix version of the own- and cross-partial derivatives, and provide a computational approach to calculating these. Extending their approach to the case of our static panel, let $d(\cdot)$ represent the $NT \times 1$ vector on the diagonal of a diagonal matrix $D(\cdot)$, where the non-diagonal elements are zeros. By construction, $D(\cdot)$ is symmetric. The $NT \times 1$ vector $d(f(\eta))$ contains the pdf (probability density function) evaluated at the predictions for each of the NT observations and associated $NT \times NT$ diagonal matrix $D(f(\eta))$ which has $d(f(\eta))$ on the diagonal. These $NT \times 1$ predictions of course depend on values taken by the explanatory variables in the $NT \times k$ matrix X that exhibit variation over both countries and time periods.

Using the matrix of own- and cross-partial derivatives, an $NT \times 1$ vector of (cumulative) *total* effects can be written as:

$$
\begin{aligned}
\left(\frac{\partial \Pr(y = 1)}{\partial x_v'}\right)\iota_{NT} &= [D(f(\eta))\iota_{NT} + \rho D(f(\eta))W\iota_{NT} + \rho^2 D(f(\eta))W^2\iota_{NT} + \ldots](I_{NT}\beta_v + (I_T \otimes W)\theta_v)\\
&= [D(f(\eta))\iota_{NT} + \rho D(f(\eta))\iota_{NT} + \rho^2 D(f(\eta))\iota_{NT} + \ldots](I_{NT}\beta_v + (I_T \otimes W)\theta_v)\\
&= (D(f(\eta))\iota_{NT})(1-\rho)^{-1}(\beta_v + (I_T \otimes W)\theta_v)\\
&= (d(f(\eta)))\,(1-\rho)^{-1}(I_{NT}\beta_v + (I_T \otimes W)\theta_v)
\end{aligned}
$$

As a scalar summary measure of *average total effect*, LeSage et al. (2011) use an average of the vector of (cumulative) *total effects* shown in (**??**).

$$
NT^{-1}(d(f(\eta))'\iota_{NT}(1-\rho)^{-1}(\beta_v + (I_T \otimes W)\theta_v) \qquad (12)
$$

To summarize the *average direct effect* LeSage et al. (2011) propose use of (**??**), where we note that $(I_T \otimes W)\theta_v = \theta_v$ by virtue of the row-stochastic nature of the matrix W.

$$\begin{aligned} \frac{1}{NT}\mathrm{tr}\left(\frac{\partial \Pr(y=1)}{\partial x_v'}\right) &= [\mathrm{tr}(D(f(\eta))) + \rho \mathrm{tr}(D(f(\eta))(I_T \otimes W) \\ &+ \rho^2 \mathrm{tr}(D(f(\eta))(I_T \otimes W^2) + \ldots] \frac{(\beta_v + \theta_v)}{NT} \end{aligned} \tag{13}$$

For the (cumulative) *average spillover effect* they propose using the difference between the average *total* effect (minus) the average *direct* effect. LeSage and Pace (2009, Chapter 4) provide several approaches to efficiently computing $\mathrm{tr}(D(f(\eta))W^j)$ which is needed to calculate the scalar summary measures for the marginal effects.

Lacombe and LeSage (2018) make a number of points regarding the scalar summary measures for *direct*, *indirect* and *total* effects estimates associated with the spatial Durbin probit model that were proposed in LeSage et al. (2011).[12] One point regards the representativeness of the scalar summary measures in (**??**) and (**??**). We note that changes in levels of the explanatory variables of the model over countries and time periods is embedded in $(d(f(\eta)))$, given the definition of $\eta = S(\rho)Z\delta$, and due to the non-linearity of the probability response in the dependent variable to these changes reflected in the pdf $f(\eta)$.

In conventional probit regressions a common way to explore the non-linearity in this relationship is to calculate "marginal effects" estimates using particular values of the explanatory variables (e.g., mean values or values constructed from quintile intervals used in place of $Z\delta$). The motivation for this practice is consideration of how the impact of changing explanatory variable values varies across the range of values encompassed by the sample data. Given the non-linear nature of the normal cumulative density function (CDF) on which the conventional probit model relies, we know that changes in explanatory variable values near the mean will have a very different impact on decision probabilities than changes in very low or high values of the explanatory variables. Lacombe and LeSage (2018) point out that the non-linear nature of the partial derivative relationship is exacerbated by the presence of another non-linear function, $S(\rho) = (I_{NT} - \rho(I_T \otimes W))^{-1} = I_{NT} + \rho(I_T \otimes W) + \rho^2(I_T \otimes W^2) + \ldots$ in the spatial probit model.

[12] Their cross-sectional spatial Durbin probit model is analogous to our decision-dependence panel probit model where the matrix W is defined on the basis of spatial proximity of countries.

As an example related to adoption of cash transfer programs in our application, consider the impact of the population size characteristic of countries (used as an explanatory variable) on the probability of adopting the program. Evaluating the partial derivative impacts for a very small population country might produce a very different probability response relative to these impacts for a very large population country. To conform with standard probit assumptions that the disturbances have a constant scalar error variance: $\sigma_\varepsilon^2 = 1$, explanatory variables are transformed to deviations from means. Nonetheless, countries with population far below the mean versus those with population far above the mean could give rise to large (non-linear) differences in the $(d(f(\eta)))$.

For this reason, we calculated effects based on the entire $NT \times 1$ vector $z^v \delta$, for each explanatory variable v, and then averaged across the T time periods to produce country-specific effects estimates. These estimates should capture differences that arise due to differences in characteristics of the sample of countries considered. In presentation of our estimation results, we show that country-specific estimates showed no significant deviation from simpler scalar summary estimates that average over all countries. This allows us to simplify presentation of the *direct*, *indirect* and *total* effects estimates, since we can ignore the large number of country-specific estimates and focus on scalar summary estimates.

4 Data and Results

We use data from 17 Latin American countries (Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, the Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, and Uruguay) and while the cash transfer data comes from actual expenditures, it uses budgeted figures when actual expenditures are missing. We also extrapolate cash transfer figures when less than two years are missing to accommodate for countries that report these figures biannually, using the simple average of the reported years. Although the data measures expenditures on these programs as a percentage of GDP, we are interested only in a dichotomous measure of the absence/presence of such programs. Specifically, for each year that a country has the cash transfer program in place, we use a 1-state variable, and for years when the program is not in place a 0-state variable.

The data for the explanatory variables come from the World Development Indicators database, in yearly frequency, and are lagged one year to allow the governments considering adoption of the program to respond to previous period conditions.[13] We use the population of each country to account for the size of the economy, the poverty rate (those living with less than 4 dollars a day) to measure the magnitude of the disadvantaged population in each country, the real GDP per capita in purchasing power parity terms to measure the level of economic development of each country, and Government Expenditures as a percentage of GDP to account for the size of the public sector in each country. Our spillover effects panel probit model allows decisions made regarding program adoption to depend not only on own-country characteristics, but also on those of other dependent countries, which are reflected by the explanatory variables matrix WX, and associated parameters θ.

4.1 Estimates of the underlying model parameters

As already noted, we use the spatial Durbin probit model to estimate the impact of our explanatory variables on the probability of countries adopting cash transfer programs. Table **??** presents the results for the spatial lag of X (SLX) and SDM probit models for both type of programs when the set j of dependent countries are defined based on out-migration flows. The SDM model subsumes the SLX specification as a special case where dependence of the adoption decision of the typical country on other countries adoption decisions is not present. (This is the case when the coefficient estimate for ρ is not significantly different from zero). Estimates for the SLX specification can be shown to be biased and inconsistent in the presence of significant dependence of adoption decisions, or in other words when the scalar coefficient ρ is different from zero. Coefficients on the X-variables in the SLX model are interpreted as direct (own-partial derivatives) effects, and those on the WX variables as reflecting spillover effects. Spillover effects in the SLX model impact only countries in the dependence set j of country i, having no higher-order dependence impacts.

The motivation for presentation of SLX probit estimates alongside those for the SDM probit specification is a test for the presence of simultaneous interaction between country adoption decisions. We note that SLX estimates can be theoretically shown to be biased and inconsistent in the face of dependence in country adoption decisions, so a comparison of estimates from these two specifications provides an indication of the relative importance of

[13] The typical endogeneity concern does not apply here because it takes time for poverty conditions to lead to program implementation, which in turn would take additional time to affect poverty - there is no contemporaneous effect between poverty and cash transfer programs. In addition, because the dependent variable reflects a binary state transition at a discrete point in time and the explanatory variable is continuous, the conventional reverse causality scenario is not likely to occur.

appropriately modeling interaction in country decisions. We see that in the case of the CCT programs, the estimate of ρ is significantly different from zero, suggesting dependence of countries adoption decisions, where the dependence set of countries are defined as countries where large shares of out-migrants from each country have settled.

The SLX coefficients reported here show that population and GDP per capita exert a statistically significant *direct* effect on CCT program adoption, and these explanatory variables together with poverty have statistically significant *spillover* effects on program adoption. This means that most characteristics of countries in the dependence set are significant in framing the context in which policy-makers operate, and are important in explaining variation in the 0,1 dependent variable values used in the SLX regression. Of course, these regression estimates are biased due to the presence of dependence between countries adoption decisions. Coefficients associated with the X and WX variables are presented for the SDM model, but as already discussed these cannot be interpreted in the usual way that regression coefficients are treated, as partial derivatives. We present *direct* and *indirect* effects estimates for the SDM probit model in the next section.

Table 2: Coefficient estimates for the spillovers probit model

	Conditional Cash Transfers		Unconditional Cash Transfers	
	SLX	SDM	SLX	SDM
Variable	coefficient	coefficient	coefficient	coefficient
Poverty	-0.141	0.150	-0.688***	-0.746***
log(population)	0.819***	1.508***	0.615***	0.659***
log(GDP per capita)	0.445**	1.001***	0.383***	0.444**
log(Gov as % GDP)	0.142	0.131	0.881***	0.979***
W*Poverty	-1.269***	-1.799***	-1.006***	-1.089***
W*log(population)	0.465***	-0.091	1.781***	1.942***
W*log(GDP per capita)	-1.422***	-2.638***	-1.255***	-1.440***
W*log(Gov as % GDP)	0.189	0.215	0.145	0.104
ρ		0.552***		0.069
Nobs, Nvars	306,25	306,25	306,25	306,25
#0,1 y-values	73,233	73,233	110,196	110,196

* indicates 90 percent significance level, ** 95 percent and *** 99 percent

For UCT programs, the SLX coefficients show that poverty exerts a negative and statistically significant *direct* effect on program adoption, while the population size, GDP per capita, and the relevance of the public sector exert a positive and statistically significant effect on program adoption. Only Government Spending as a percentage of GDP fails to exert a statistically significant *indirect* impact on program adoption. The SDM estimate for the dependence parameter ρ is found to be statistically insignificant for the UCT programs, indicating that there is no dependence between country adoption decisions in this case.

4.2 CCT program adoption results

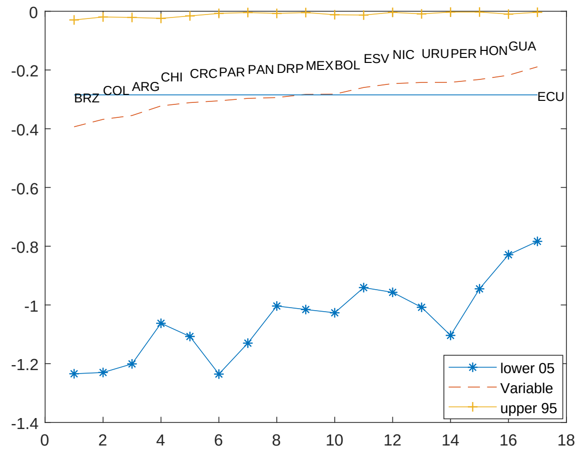
Based on underlying parameters from the SDM model specification, we calculated country-specific estimates of the *direct* and *spillover* effects that mimic conventional probit model marginal effects. Estimates for these were calculated based on the entire $NT \times 1$ vector $z^v\delta$, for each explanatory variable v, and then averaged across the T time periods to produce country-specific effects estimates for the weight matrix based on out-migration flows.

An examination of the county-specific effects estimates showed no statistically significant difference between scalar summary effects estimates based on an average across all countries and the effects for individual countries. The effects for individual countries exhibit wide confidence intervals for all variables, meaning that distinguishing between country-specific effects was not meaningful. Figure **??** provides an illustration of this point, showing country-specific estimates of the *total* effect associated with the poverty variable. The lower 0.05 and upper 0.95 credible intervals were calculated from the set of retained MCMC draws. The scalar summary measure that averages over all countries is shown as the horizonal line in the figure. Given the lower 0.05 and upper 0.95 intervals in the figure, it should be clear that despite the small differences in country-specific estimates, there are no significant differences in these estimates between countries. A similar result was found for all other *direct*, *indirect* and *total* effects estimates, where the wide lower and upper credible intervals pointed to a lack of significant differences between country-specific estimates.

This also means that the scalar summary estimate (shown as the horizon line in the figure) provides a valid basis for inference regarding the *total* effect of the poverty variable on the probability of program adoption. Recall that the concern raised by Lacombe and LeSage (2018) was that scalar summary estimates might not provide an adequate basis for inference in situations where a great deal of observation-specific variation in the effects estimates exists. This result allows us to greatly simplify presentation of the estimation results by focusing on scalar summary estimates of the *direct*, *indirect* and *total* effects estimates that reflect an average over the entire sample of countries.

Table **??** presents the country-specific effects estimates using the minimum, median and maximum effect estimate from the set of 17 country-specific estimates produced. Despite the variation in magnitudes shown in the table, as already noted, these differences in magnitude

Figure 3: Observation-level Total Effect for Poverty



are not statistically significant.

Turning first to the *direct* effects of CCT program adoption, the results indicate that countries with lower poverty rates have a higher probability of CCT program adoption, and this *direct* effect is statistically significant for all countries based on lower 0.05 and upper 0.95 credible intervals constructed using the MCMC draws.[14] These results suggests that lower poverty rates have a positive impact on the probability of program adoption, which could mean that countries with lower poverty maintain (or expand) CCT programs in an effort to be perceived as the generators of these trends, and to receive future political support.[15] It is also possible that countries with levels of poverty beyond a certain threshold level fear adverse budgetary impacts that would arise from high participation rates in these programs, whereas countries where poverty is trending downward have reached a poverty threshold that eliminates these concerns. A reviewer suggested possible endogeneity of the poverty rate variable. One way to test for endogeneity is to regress the vector of N countries future period poverty on the N countries past period program adoption states, e.g, $pov_{t+1} = \alpha + \beta y_t + \varepsilon_t$. If poverty is not endogenous with regard to program adoption states, then the coefficient estimate for β will be zero. The results from this regression for each year showed no cases where the null-hypothesis is rejected at the 99% level (and only

[14] A sequence of 2,500 MCMC draws were used with the first 500 discarded for burn-in of the sampler.

[15] We also examined World Bank measures for extreme poverty that use a 1.9 dollars per day and a regional measure of 2.5 dollars per day, as well as a Gini coefficient measure. All of these produced a negative *direct* effect estimate.

two cases where it is rejected at the 95% level).

In terms of the impact of *population* on the probability of CCT program adoption, we find that larger countries have a higher probability of adoption, and the *direct* effects are significant for all countries. Intuitively, the number of people experiencing harsh conditions are more visible in large countries, which should exert pressure on their governments to institute redistributive programs. Our measure of development, *GDP per capita*, is also found to exert a negative and statistically significant *direct* effect on CCT program adoption, indicating that countries with lower levels of development have a higher probability of program adoption. This effect can be rationalized as an extension of the welfare state in poorer countries, where there is more dependence on the public sector to solve welfare issues. Our last explanatory variable, the relative size of the government sector in the economy does not have a statistically significant effect on the probability of program adoption, suggesting that the size of the government is irrelevant to adoption of CCT redistributive programs.

Given our reliance on a weight matrix defined by out-migration flows, information that migrants from an origin country residing in a host country send to their relatives and friends back home are presumed to affect policymaking in the country of origin. A question of interest for this model specification is whether the *spillover* (or *indirect*) effects are significant or not, because significant *spillover* effects would indicate that a model specification that ignores other countries' influences on the probability of program adoption would produce estimates that are biased and inconsistent. In other words, a model that ignores country decision interaction would set *spillover* effects to zero (restricting the parameter ρ to zero), leaving only *direct* effects, which would be biased. Table **??** shows the *indirect* effect estimates for the four explanatory variables, where we see that three of the four variables exhibit significant *indirect* effects in the case of the CCT programs. The signs of these *indirect* effects are the same as those of the *direct* effects presented in the Table, compounding the effect of a particular measure on the probability of adopting these CCT programs. The magnitudes reported in the table for *indirect* or *spillover* effects indicate that these account for a substantial portion of the *total* effects arising from changes in the explanatory variables on the probability of program adoption. Specifically, we find magnitudes similar to those of the *direct* effects, for example, the *Poverty* variable has *direct* effects ranging from -0.088 to -0.184, with *indirect* effects ranging from -0.083 to -0.172.

Table 3: Summary of effects estimates for the probability of adopting cash transfer programs

	Conditional Cash Transfers (CCT)	Unconditional Cash Transfers (UCT)
Direct Effects		
Poverty		
Minimum	-0.1844*	-0.3484*
Median	-0.1412*	-0.2281*
Maximum	-0.0888*	-0.1214*
Population		
Minimum	0.0943*	0.2016*
Median	0.1485*	0.3774*
Maximum	0.1934*	0.5779*
GDP per capita		
Minimum	-0.1674*	-0.1546*
Median	-0.1262*	-0.0995*
Maximum	-0.0800*	-0.0498*
Government spending		
Minimum	0.0308	0.0897*
Median	0.0491	0.1734*
Maximum	0.0653	0.2674*
Indirect Effects		
Poverty		
Minimum	-0.2357*	-0.0210
Median	-0.1989*	-0.0128
Maximum	-0.0903*	-0.0055
Population		
Minimum	0.0636*	0.0092
Median	0.1394*	0.0215
Maximum	0.1643*	0.0351
GDP per capita		
Minimum	-0.2247*	-0.0085
Median	-0.1885*	-0.0051
Maximum	-0.0864*	-0.0022
Government spending		
Minimum	-0.0255	0.0042
Median	-0.0207	0.0098
Maximum	-0.0086	0.0163
Total Effects		
Poverty		
Minimum	-0.3634*	-0.3710*
Median	-0.2853*	-0.2426*
Maximum	-0.1774*	-0.1295*
Population		
Minimum	0.1840*	0.2151*
Median	0.2950*	0.4022*
Maximum	0.3753*	0.6137*
GDP per capita		
Minimum	-0.3323*	-0.1659*
Median	-0.2578*	-0.1065*
Maximum	-0.1600*	-0.0533*
Government spending		
Minimum	0.0614	0.0959*
Median	0.0984	0.1853*
Maximum	0.1315	0.2852*

* indicates 90 percent significance level, ** 95 percent and *** 99 percent

The negative indirect effect for *poverty* rates suggests that countries operating in a contextual setting where the set of dependence countries have lower poverty rates are more likely to adopt CCT programs, whereas countries whose dependence set consists of countries with higher poverty rates are less likely to adopt CCT programs. Since our dependence set of countries is based on countries where a large proportion of out-migrants from country i reside, this implies that a typical country i having ex-patriots residing in lower poverty rate countries is more likely to adopt CCT programs, and a typical country i having ex-patriots residing in higher poverty rate countries less likely to adopt CCT programs. This seems consistent with the notion that countries whose out-migrants are moving to lower poverty rate countries may be engaged in competitive attempts to entice out-migrants to stay in the home country rather than seek economic opportunity in countries with less poverty.

The negative *indirect* effect found for our *GDP per capita* measure suggests an inverse relationship between adoption of transfer programs and income of the set of j countries on which adoption decisions are dependent. This suggests that countries operating in a contextual setting where the set of dependence countries have lower *GDP per capita* are more likely to adopt CCT programs, whereas countries whose dependence set consists of countries with higher *GDP per capita* are less likely to adopt CCT programs. Since our dependence set of countries is based on countries where a large proportion of out-migrants from country i reside, this implies that a typical country i having ex-patriots residing in lower *GDP per capita* countries are more likely to adopt CCT programs, and a typical country i having ex-patriots residing in higher *GDP per capita* countries less likely to adopt CCT programs. Here again, the fact that out-migrants are moving to lower *GDP per capita* countries might make it appear that home countries have a chance at retaining out-migrants through implementation of CCT programs.

The *positive* indirect effect for *Population* suggests that countries operating in a contextual setting where the set of dependence countries are larger in population size are more likely to adopt CCT programs, whereas countries whose dependence set consists of countries with smaller population are less likely to adopt CCT programs. Since our dependence set of countries is based on countries where a large proportion of out-migrants from country i reside, this implies that a typical country i having ex-patriots residing in large population size countries are more likely to adopt CCT programs, and a typical country i having ex-patriots residing in smaller population sized countries less likely to adopt CCT programs.

This result seems consistent with the fact that larger population sized countries such as Brazil and Mexico have successful cash transfer programs in place, and information transfer from these countries to the home country provide political pressure for implementing CCT programs.

Since the sum of the two effects (*direct* and *indirect*) represent the (cumulative) *total* effect associated with a change the explanatory variables, and since both sets of estimates present the same sign, the *total* effects of our four explanatory variables are the same as the *direct* and *indirect* effects discussed above. For example, a lower own-country *Poverty* increases the probability of CCT program adoption, and also increases the probability of CCT adoption in the set j of dependent countries, those countries where a large number of ex-patriots reside.

In the case of the UCT programs, presented in the right column of Table **??**, we find that the *direct* effect estimates associated with the *Poverty*, *Population*, and *GDP per capita* measures are similar in signs and significance to those found for the CCT program, but *Government spending* also becomes statistically significant for the UCT programs. Lower *Poverty* and *GDP per capita* increase the probability of cash transfer program adoption, while lower *Population* reduces the probability of adopting UCT programs in Latin America. In terms of *Government spending*, we find that it exerts a positive and significant *direct* effect on UCT program adoption over time, suggesting that countries with larger public sectors have a higher probability of adopting UCT programs. Having the resources, or the economic structure with larger public sectors, facilitates the adoption of these types of programs, as might be expected.

Turning to the *indirect* effects estimates for the UCT programs, we find that the country-level *indirect* effects are not statistically significant for this type of program. This suggests that implementation of the UCT programs by a typical country i does not exert a significant impact on other countries in the dependence set. This was true for all three types of dependence sets explored here. The estimated ρ for the case of a spatial dependence set was 0.133, for the in-migration dependence set -0.007, and for the out-migration dependence set 0.069, and none of these were statistically significant based on the lower 0.05 and upper 0.95 credible intervals. These results are reported in the supplemental material available in the author's webpage.

The insignificant *indirect* effect estimates results in country-level *total* effects estimates that are similar in magnitude to the *direct* effects, since these are the sum of *direct* plus *indirect* effects estimates, and present the same statistical significance as the *direct* effects. The *total* effects magnitudes are slightly larger than the CCT estimates found for *Population* and *Government spending*, but are slightly smaller than the CCT estimates found for *Poverty* and *GDP per capita*. Table **??** is consistent with these findings, since the SLX and SDM estimates are similar both in terms of magnitude and significance, since dependence of the adoption decision for the typical country on other countries adoption decisions is not present.

5 Conclusion

Public cash transfer programs have been implemented to reduce overall income inequality and poverty rates, but have been shown to promote long-term growth as well, because they tend to improve human capital. Universal, or unconditional cash transfers (UCT) are less redistributive in nature, and have a lower contribution to human capital, but are significantly larger in Latin America (even if they cover fewer people). Conditional cash transfers (CCT), on the other hand, are smaller in size but reach more people, and are better targeted to those in need, improving their standard of living through their contribution to educational attainment and health care access, boosting GDP per capita in the long term. These programs have gained prominence throughout Latin America for their perceived impact on reducing poverty and improving the reach of education, and are being introduced in many other countries around the world because of the perceived benefits.

In contrast to past literature regarding these programs, we examine the factors that lead to the adoption of these programs in 17 Latin American countries, and test for the presence of significant spillover effects arising from dependence of program adoption decisions made by each country on a set of j other countries. The empirical results are consistent with significant *spillover* impacts for the CCT programs, and this conclusion is robust to varying definitions of connectivity structures used to model dependence between countries. This implies that decisions to adopt conditional cash transfers programs are significantly influenced by the presence/absence of these programs other countries.

In terms of own-country characteristics that significantly impact the probability of cash transfers programs, we find that larger countries (measured by population) had a higher probability of adopting these programs over time, in line with the conventional wisdom. *Poverty* is found to have a significant but negative impact on the probability of program adoption (for a number of different poverty measures). The somewhat counterintuitive negative impact might arise because countries with poverty beyond some threshold level may be reluctant to adopt these programs because of budgetary concerns that would arise in the face of widespread program participation. It might also be the case that in countries where poverty begins trending downward over time, political leaders want to be perceived as responsible for these downward trends as a result of their implementation of popular cash transfer programs.

The nature of *indirect* effects differs between CCT and UCT programs, with the former displaying positive *indirect* effects and the later insignificant *indirect* effects. Positive *indirect* effects for the CCT programs mean that the presence (absence) of a program in the typical country is – positively – associated with the presence (absence) of the program in the group of countries on which program adoption decisions depend. Presumably, sharing of information regarding positive experiences with outcomes such as educational attainment and health care access between connected countries leads to this type of positive association. The insignificant *indirect* effects found for UCT programs mean that adoption decisions are made independently of the decisions taken by other countries, so we see no systematic pattern of adoption or lack thereof between a typical country and other countries in our sample of 17 Latin American countries.

Lastly, we find that a specification where dependence is defined based on the relative importance of ex-patriots from a particular country living in a host country is most consistent with our sample data, based on log-probability and quadratic probability scores (LPS, QPS). This seems intuitively plausible and suggests that information regarding these programs in the destination countries is transmitted by ex-patriots to their relatives back home. Alternative information transmission channels based on spatially neighboring countries and those from which a large proportion of in-migrants arrived were found to produce slightly lower LPS, QPS. Nonetheless, we note that dependence defined based on spatial proximity, in- and out-migration patterns were highly correlated and produced similar estimates and inferences.

References

Aizer A, Eli S, Ferrie J, Lleras-Muney A (2016) The Long-run Impact of Cash Transfers to Poor Families. *American Economic Review*, 106(4): 935-971.

Albert J, Chib S (1993) Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, 88: 669-679.

Behrman JR, Parker SW, Todd PE (2011) Do Conditional Cash Transfers for Schooling Generate Lasting Benefits? A Five-Year Follow up of PROGRESA/Oportunidades. *Journal of Human Resources*, 46(1): 93-122.

Behrman JR, Sengupta P, Todd PE (2005) Progressing through PROGRESA: An impact Assessment of a School Subsidy Experiment in Rural Mexico. *Economic Development and Cultural Change*, 54(1): 237-275.

Bianchi M, Bobba M (2013) Liquidity, Risk, and Occupational Choices. *Review of Economic Studies*, 80(2): 491-511.

CEPAL (2019) Non-contributory Social Protection Programmes in Latin America and the Caribbean Database. *CEPAL*.

Cespedes N (2014) General Equilibrium Analysis of Conditional Cash Transfers. *Peruvian Economic Association Working Paper No. 25*.

Elhorst JP, Heijnen P, Samarina H, Jacobs JPAM (2017) State Transfers at Different Moments in Time: A Spatial Probit Approach. *Journal of Applied Econometrics* 32(2): 422-439.

Gertler PJ, Martinez SW, Rubio-Codina M (2012) Investing Cash Transfers to Raise Long-Term Living Standards. *American Economic Journal: Applied Economics*, 4(1): 164-192.

Lacombe DJ, LeSage JP (2018) Use and interpretation of spatial autoregressive probit models. *Annals of Regional Science*, 60(1): 1-24.

LeSage JP, Pace RK, Lam N, Campanella R, Liu X (2011) New Orleans business recovery in the aftermath of Hurricane Katrina. *Journal of the Royal Statistical Society: Series A* 174(4): 1007-1027.

LeSage JP, Pace RK (2009) *Introduction to spatial econometrics.* Boca Raton: CRC Press.

LeSage JP, Pace RK (2014) The biggest myth in spatial econometrics, *Econometrics* 2: 217–249.

Parker SW, Todd PE (2017) Conditional Cash Transfers: The Case of Progresa/Oportunidades. *Journal of Economic Literature*, 55(3): 866-915.

Parker SW, Vogl T (2018) Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico. *NBER Working Paper* No. 24303.

Peruffo M, Cavalcanti Ferreira P (2017) The Long-Term Effects of Conditional Cash Transfers on Child Labor and School Enrollment. *Economic Inquiry*, 55(4): 2008-2013.

Robles M, Rubio MG, Stampini M (2015) Have Cash Transfers Succeeded in Reaching the Poor in Latin America and the Caribbean? *Inter-American Development Bank Working Paper.*

Schultz TP (2004) School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program. *Journal of Development Economics*, 74(1): 199-250.

Todd PE, Wolpin KI (2006) Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility. *American Economic Review*, 96(5): 1384-1417.

Vacaflores DE (2019) The Effectiveness of Governmental Cash Transfers in Promoting Economic Growth and Reducing Income Inequality. *Texas State University Working Paper.*