The Welfare Effects of Encouraging Rural-Urban Migration

David Lagakos
UCSD and NBER

Mushfiq Mobarak
Yale University and NBER

Michael E. Waugh
New York University and NBER

March 2017

ABSTRACT

This paper studies the welfare effects of encouraging rural-urban migration in the developing world. To do so, we build a dynamic incomplete-markets model of migration in which heterogenous agents face seasonal income fluctuations, stochastic income shocks, and disutility of migration that depends on past migration experience. We calibrate the model to replicate a field experiment that subsidized migration in rural Bangladesh, leading to significant increases in both migration rates and in consumption for induced migrants. The model's welfare predictions for migration subsidies are driven by two main features of the model and data: first, induced migrants tend to be negatively selected on income and assets; second, the model's non-monetary disutility of migration is substantial, which we validate using using newly collected survey data from this same experimental sample. The average welfare gains are similar in magnitude to those obtained from an unconditional cash transfer, though migration subsidies lead to larger gains for the poorest households, which have the greatest propensity to migrate.

Email: lagakos@ucsd.edu, ahmed.mobarak@yale.edu, mwaugh@stern.nyu.edu. For helpful comments we thank Greg Kaplan, Louis Kaplow, Sam Kortum, Melanie Morten, Paul Niehaus, Natalia Ramondo, Chris Tonetti and seminar participants at Edinburgh, the Einaudi Institute, Fordham, the Hong Kong School of Economics and Finance, NYU, St. Andrews, Stockholm IIES, UC Irvine, USC Marshall, Washington, Yale, the Minnesota Macro Workshop, the MadMac Growth conference, the NBER Macroeconomics Across Time and Space Meeting, the SED meeting and the AEA meeting. For outstanding research assistance we thank Elizabeth Carls, Menaal Ebrahim, Patrick Kiernan and Seungmin Lee, and for financial support we thank the International Growth Centre. All potential errors are our own.

1. Introduction

Differences in income per capita across countries are accounted for in large part by differences in total-factor productivity (TFP) (see e.g. Hall and Jones, 1999; Caselli, 2005). Following the seminal work of Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), much recent emphasis in the macroeconomics literature has been placed on resource misallocation as a theory of TFP. The idea is that aggregate efficiency of production will be lower when factors of production are badly allocated across firms, sectors or regions within an economy.[1]

One potentially large source of misallocation in developing countries is the distribution of workers across space, and in particular across rural and urban areas (Restuccia, Yang, and Zhu, 2008; Vollrath, 2009; McMillan and Rodrik, 2011; Hnatkovska and Lahiri, 2013; Bryan and Morten, 2015). Most developing countries have large shares of the population living in rural areas and working in agriculture. Yet in almost every developing country, there are large gaps in average income per head between rural and urban areas, and relatedly, between agricultural and non-agricultural workers (Young, 2013; Gollin, Lagakos, and Waugh, 2014). If these gaps reflect misallocation, then encouraging workers to move out of less-productive rural areas could yield substantial welfare gains.

Recent experimental evidence suggests that rural workers may indeed be misallocated, at least to some extent. Bryan, Chowdhury, and Mobarak (2014) show that, in Bangladesh, relatively small subsidies to migration result in large increases in seasonal migration and consumption for those households induced to send a migrant. Other longitudinal data, though not experimental in nature, also show large increases in consumption for migrants. For example, Beegle, De Weerdt, and Dercon (2011) use detailed tracking surveys from Tanzania to show that the household group with the highest average consumption increases over the 1990s were those that moved from rural to urban locations.

An opposing perspective on these urban-rural gaps is that they reflect efficient outcomes, driven by disutility of migration, broadly defined, or sorting of more productive workers to urban areas (Lagakos and Waugh, 2013; Young, 2013; Herrendorf and Schoellman, 2016). Evidence shows that workers in urban areas – and rural-urban migrants – tend on average to be those with more education (Young, 2013), higher returns to schooling (Herrendorf and Schoellman, 2016) and higher scores on tests of cognitive ability (Miguel and Hamory, 2009). Migration is also costly in non-monetary terms, so that even if migration can raise

[1] Channels for misallocation emphasized in the recent literature include financial frictions (Buera and Shin, 2013; Midrigan and Xu, 2014; Moll, 2014), information frictions (David, Hopenhayn, and Venkateswaran, 2016), adjustment costs (Asker, Collard-Wexler, and De Loecker, 2015), heterogenous markups (Peters, 2016), entry frictions (Yang, 2016) and delegation frictions (Akcigit, Alp, and Peters, 2016), size-dependent policies (Guner and Xu, 2008), regional differences in tax rates (Fajgelbaum, Morales, Serrato, and Zidar, 2015), among others.

consumption, it may be partially offset by poor urban housing options for migrants, dislike of being away from family members, or less access to village risk-sharing networks (Harris and Todaro, 1970; Morten, 2013; Brueckner and Lall, 2015; Munshi and Rosenzweig, 2016).

The size of the welfare gains from encouraging internal migration is important from a policy perspective, and many developing countries take opposing policy oppositions. China famously prevents rural-urban migration through its Hukou system, while others implicitly discourage migration through rural employment guarantee programs that pay households to remain in rural areas to work there (Beegle, Galasso, and Goldberg, 2016; Muralidharan, Niehaus, and Sukhtankar, 2016). On the other side, some countries, such as Bangladesh and Indonesia, are experimenting with conditional migration transfer programs that encourage rural-urban migration during unproductive agricultural seasons.

In this paper, we study the welfare effects of policies that encourage rural-urban migration in developing countries. To do so, we build a dynamic model of migration in which heterogenous households choose to locate in either an urban region or rural region. Households are heterogenous in their degree of permanent productivity advantage in the urban area, as in Roy (1951). In addition, agents face deterministic seasonal income fluctuations and stochastic income shocks. Markets are incomplete, and agents insure themselves through a buffer stock of savings as in Bewley (1977), Aiyagari (1994) and Huggett (1996), and following a large literature in macroeconomics (see e.g. Heathcote, Storesletten, and Violante, 2009; Kaplan and Violante, 2010). Households can migrate either temporarily or permanently across locations, and face disutility of migration that depends on past migration experiences. In particular, households that have recently migrated face a chance of migration without cost in subsequent periods. This is meant to capture the connections that a migrant may make with an employer or landlord in the urban area, say, which improve the prospects of future migration.

While the model abstracts from reality in many ways, it is rich enough to capture multiple forces relevant for the migration decisions. In particular, it allows for permanent sorting on permanent comparative advantage as in Lagakos and Waugh (2013) and Young (2013). It allows for temporary migration due to transitory shocks, as in Kennan and Walker (2011), or due to predictable seasonal downturns, as emphasized by Khandker (2012). It allows the degree of income risk to vary across regions, including higher risk in urban areas, as in the theory of Harris and Todaro (1970). Finally, the model allows for repeat migrants to behave differently from first-time migrants, even for agents who are otherwise observationally similar, which is consistent with evidence on repeat migration episodes. Importantly, the model allows for workers to be misallocated across locations due to incomplete financial markets. The extent of workers misallocation depends on parameter values, and more generally on the data used to calibrate the model. Compared to previous dynamic migration models, ours allows for less richness in the number of locations than do Kennan and Walker (2011) or Caliendo, Dvorkin, and Parro (2015). On the other hand, our model allows for migration decisions that depend jointly on shocks, the stock of savings and past migration experience.

Our paper departs from the previous migration literature mostly in how it disciplines the model quantitatively. In particular, we replicate the results of a randomized controlled trial (RCT) within the model, and use simulated method of moments to match the model's outcomes to the experimental data. The RCT we replicate is that of Bryan, Chowdhury, and Mobarak (2014), mentioned above, which provided a modest subsidy to seasonal migration in rural Bangladesh, leading to substantial increases in migration and consumption for those induced to migrate. The main moments of the experiment that we target are: (i) the increase in seasonal migration rate resulting from the subsidy, which was 22 percent, (ii) the consumption increase for those induced to migrate, which was 30 percent, and (iii) the increase in seasonal migration a year hence, when the subsidy was no longer offered, which was 9 percent. We also match a number of cross-sectional moments from a nationally representative household survey from Bangladesh, including the rural-urban wage gap and variances of log wages.

Matching these experimental and cross-sectional moments helps discipline the characteristics of workers that are near the margin, relative to those that are regularly migrating already or permanently located in cities. To match the data, the model requires that workers near the margin are negatively selected on productivity and assets. The reason is that the model matches the low OLS regression coefficient of migration on consumption and the high experimental local average treatment effect (LATE) of migration on consumption. The model can only match these moments if workers that are induced to migrate have relatively low productivity levels. In addition, the calibrated model requires a large disutility of rural-urban migration and an effect of past experience on current migration utility that is highly transitory. In other words, migration disutility must be high for new migrants, and migration experience, which reduces migration disutility, must be easy to lose for previous migrants. The reason is that the model matches the large consumption gain of induced migrants but the low likelihood of repeat migration for induced migration multiple years hence.

We then use the calibrated model to quantify the welfare gains from subsidizing rural-urban migration. We find that while migration subsidies raise consumption substantially, they raise welfare by a more modest amount. The reason is that, in the model, the consumption gains for migration are offset in large part by the disutility of migration and high probability of losing migration experience. On average, households offered a conditional migration transfer equal to around two-weeks wages gain 0.2 percent in consumption-equivalent welfare. Gains were the highest among the households with the lowest income and assets, who were most likely to migrate in response to the offer, and on the order of 2.0 percent in consumption-equivalent welfare. As a frame of reference, an unconditional transfer program costing the same total amount raises welfare of the the poorest households by 1.5 percent in consumption equivalents, while the richest households gain more from the unconditional transfer (since the majority are not induced to migrate by the conditional transfer).

Given that our model produces such large migration disutility, which reduces the welfare gains from migration, we are left wondering what exactly this migration disutility is. To get some insight on this question, and validate the fact that this disutility is in the data, we conduct a new discrete-choice experiment on the same population used to calibrate the model, using surveys about hypothetical migration options. In the survey, respondents were asked to choose between a discrete set of migration opportunities varying in employment risk, wages conditional on employment, housing options at destination, and frequency of return visits. Using the survey responses, we estimate a multinomial logit model to compute the marginal effects of each migration characteristic on migration. The estimates point to substantial disutility of migration, in particular for the poor housing options in urban areas, which reduce migration rates by 17.4 percent all else equal.This is consistent with our model's prediction of a substantial disutility of migration.

We conclude that the extent of worker misallocation between rural and urban areas is likely to be modest, at least in Bangladesh. Our welfare gains from encouraging migration lead to welfare gains that are substantially lower than those implied by Hsieh and Klenow (2009) for capital misallocation in Indian and Chinese manufacturing, or factor misallocation across farms found by Adamopoulos and Restuccia (2014), Restuccia and Santaeulalia-Llopis (2016) and Adamopoulos, Brandt, Leight, and Restuccia (2016). In terms of methodology, our work follows the seminal papers by Todd and Wolpin (2006) and Kaboski and Townsend (2011), which discipline dynamic structural models using quasi-experimental evidence, rather than non-experimental moments, as is most common in macroeconomics. Our paper builds on these by estimating our structural model directly using variation induced by an RCT, where concerns about endogeneity are even less present.[2] Our work also relates to McKenzie, Gibson, and Stillman (2010), who finds a very strong role for selection in explaining wage differences between international migrants and non-migrants from Tonga, and to Bazzi, Gaduh, Rothenberg, and Wong (2016), who find relatively modest gains from rural-to-rural migra-

[2]Other papers bridging the gap between general equilibrium models of development and RCTs is Buera, Kaboski, and Shin (2014), who use a macro model to help interpret the general-equilibrium effects of unconditional asset transfers studied by Bandiera, Burgess, Das, Gulesci, Rasul, and Sulaiman (Forthcoming), and Greenwood, Kircher, Santos, and Tertilt (2013), who build a general equilibrium model of the AIDS epidemic to complement the many RCT's by including behavioral responses to changing in infection probabilities.

tion in Indonesia, because skills are largely non-transferable across rural regions.[3]

2. Model of Migration

In this section, we present our model of migration. Workers are heterogenous in permanent productivity levels in the urban area, and choose their location each period given monetary and non-monetary (utility) migration costs which depend on past migration experience. Agents face deterministic seasonal income fluctuations as well as stochastic income shocks, and use a single asset to self-insure themselves. For simplicity, we focus on a stationary distribution of the model, the fraction of workers in each region and other aggregate variables remains constant each period, as does the distribution of workers by state.

2.1. Economic Environment

Preferences. Households are infinitely lived and maximize expected discounted utility

$$\sum_{t=0}^{\infty} \beta^t u(c_t)\bar{u}^{x_t} \tag{1}$$

where $u(c_t) = c_t^{1-\alpha}/(1-\alpha)$, α is the coefficient of relative risk aversion, β is the discount factor and c_t is household consumption. The variable $\bar{u}$ captures the non-monetary costs of migration, and $x_t \in \{0, 1\}$ is an indicator variable representing whether the household is an "inexperienced migrant" or not.

Inexperienced migrants experience disutility $\bar{u}$ if they locate in the urban area in period t, whereas experienced migrants experience no such disutility. After each period in the urban area, inexperienced migrants because experienced with probability $1 - \lambda$. This is meant to capture any way in which rural-urban migrants become accustomed to being in urban areas, say by developing a network of friends or potential employers. Experienced migrants can become inexperienced again by returning to the rural area. Each period in the rural area, the probability that an experienced migrant becomes inexperienced again is $1 - \mu$.[4]

The motivation behind these modeling choices are twofold. First, we want to model the fact that migrants dislike certain aspects of migrating to an urban area (see the discussion

[3]The role of risk in explaining urban-rural wage gaps dates back to Harris and Todaro (1970), who allows for an urban traditional sector, and has been followed by many others, including Rauch (1993), who links the risk of urban underemployment to overall income inequality. None of these studies have had the benefit of experimental data on migration, however, to guide their answers, as we do.

[4]This formulation is related to, but distinct form, locations being "experience goods" in migration in migration models, as in Kaplan and Schulhofer-Wohl (2017). Our model households know for certain what the migration disutility is, but that disutility may fall after they move, and remain low for some time even after returning.

in Section 4.1). However, we also want to model the idea the utility that one experiences from a location improves as one becomes accustomed to living in a location, in way that is proportional to the amount of time spent there.

Endowments. Households supply one unit of labor inelastically, with efficiency units that vary across time and across locations as in Roy (1951). Households differ in permanent productivity z in the urban area which is drawn from distribution $G(z)$, and is strictly greater than zero. Households are identical in rural permanent productivity and this value is normalized to one. Thus, the vector $\{1, z\}$ describes a household's permanent productivity in the rural and urban area.

Households experience transitory shocks to their endowments. Denoting s as the current transitory shock, this shock is represented by a finite state Markov chain with $\mathcal{P}(s'|s)$ being the transition matrix that describes the probability a household transits to state s'. To allow for this shock to have a differential impact on earnings (and risk) across locations, we assume that the household specific, transitory component on wages for the rural area is s and s^{γ} for the urban area.

The parameter γ parameterizes differentials in risk across locations. In particular, if $\gamma > 1$, this formulation will imply that shocks in the urban area have a larger impact on incomes that in the rural area. And, hence, the urban area will be riskier than the rural area. The benefit of this modeling choice is that it allows us reduce the dimensionality of the state-space to just focusing on one shock (versus multiple shock-processes across locations). Still, it captures the old idea in economics that differential risk in urban and rural areas may be a deterrent to migration as a well as a source of urban-rural average income differences (Harris and Todaro, 1970).

Production. There is one homogenous good produced in both locations. Locations differ in the technologies they operate. The rural technology is

$$Y_r = A_r^i N_r^{\alpha}, \tag{2}$$

where N_r are the effective labor units working in the rural area, $0 < \alpha$¡1 so that there is a decreasing marginal product of labor in the rural area, and A_r^i is rural productivity indexed by season i. Seasonality is modeled with the rural area experiencing deterministic, seasonal fluctuations. Specifically, rural productivity takes two values: $i \in \{g, \ell\}$ with productivity values satisfying $A_r^g > A_r^{\ell}$, where if current rural productivity is A_r^g, then the economy transits to productivity state A_r^{ℓ} next period. Superscript g is for "growing" season, and superscript

ℓ is for "lean" season. The urban technology is

$$Y_u = A_u N_u \tag{3}$$

where N_u are the effective labor units supplied by households working in the urban area. Notice that N_u and N_r do not sum to one, due to the Roy assumptions on individual productivity, and in general depend on the temporary shock realizations and the entire allocation of workers across sectors.

Wages. In season i, with N_r workers in the rural area, wages per efficiency unit are

$$\omega_{r,i}(N_r) = A_r^i \alpha N_r^{\alpha-1} \quad \text{and} \quad \omega_u = A_u. \tag{4}$$

Agents working in a particular location receives wages that are the product of (4) and the number of their efficiency units (both in permanent and transitory terms). We denote the labor income a household with with permanent state $\{1, z\}$ and transitory state s receives as for working in location i as:

$$w_r(z, s, i) = s\omega_{r,i} \quad \text{and} \quad w_u(z, s) = zs^{\gamma}\omega_u, \tag{5}$$

which depends on the product of a households permanent and transitory productivity and wages per efficiency unit described in (4).

Location Options. Households have choices about where to reside and work. Households in the rural area have three options. First, they can work in the rural area. Second, they can pay the fixed cost m_T and work in the urban area for one period and return to the rural area the next period. This is (temporary) seasonal migration in the model: a one-period working spell in the urban area by a rural household. Third, the household can pay the fixed cost $m_P > m_T$, and work in urban area for the indefinite future. This is permanent migration: a move that enables the household to permanently live and work in the urban area.

Households residing in the urban area have similar options. They can work in the urban area. Or they can pay fixed cost m_P and work in rural area for the indefinite future. This latter option allows for rural to urban and then urban to rural moves as a household's comparative advantage changes over time.

Asset Choices. Households can accumulate a non-state contingent asset, a, with gross rate of return, R. Asset holdings are restricted to be non-negative and, thus, there is no borrowing. Furthermore, we assume that R is exogenous.

2.2. Optimization

Before describing the value functions of a household, it is important to have a complete accounting of the state space. The state variables for a household can be divided into things that are permanent, transitory, endogenous and aggregate.

- **Permanent productivity state.** Each household is endowed with z efficiency units in the urban area and one efficiency in the rural area. This is the "static Roy model" aspect of the model.

- **Transitory productivity state.** Each household is subject to transitory productivity shocks, s.

- **Endogenous state variables.** There are three endogenous (individual) state variables. The first is the household's asset holdings, a. The second is a composite variable which describes the household's location and migration status. The possible states are: rural, seasonal-migrant (which means they are living in the rural area but working in the urban area for one period), and urban. The third is whether the household is an inexperienced migrant or not, x. Inexperienced migrants suffer disutility $\bar{u}$ from locating in the urban area.

- **Aggregate state variables.** There are two aggregate state variables: the season, $i \in \{g, \ell\}$, and the number of workers in the rural area, N_r. The season determines the current and future productivity in the rural area, and jointly, the two aggregate states determine the current wage per efficiency units as in equation (4).

We begin with the problem of a rural household. Because z is time-invariant for each household, we omit

Rural Households. A rural household with productivity z solves the following problem:

$$v(a, r, s, x, i, N_r) = \max \left\{ v(a, r, s, x, i, N_r, \mid \text{stay}),\; v(a, r, s, x, i, N_r, \mid \text{seas}),\; v(a, r, s, x, i, N_r, \mid \text{perm}) \right\} \quad (6)$$

where a household chooses between staying in the the rural area, seasonally moving, and permanently moving. Conditional on staying in the rural area, the value function is:

$$v(a, r, s, x, i, N_r \mid \text{stay}) = \max_{a' \in \mathcal{A}} \left[u(Ra + w_r(s, i, N_r) - a') + \beta\, \mathbb{E}[v(a', r, s', x', i', N_r')] \right] \quad (7)$$

which says the household chooses future asset holdings to maximize the expected present discounted value of utility. The asset holding must respect the borrowing constraint and, thus, must lie in the set $\mathcal{A}$. Given asset choices, a household's consumption equals the gross return on current asset holdings, Ra, plus labor income from working in the rural area, $w_r(z, s, i)$, minus future asset holdings. The next period's state variables are the new asset holdings, location in the rural area, the subsequent season, the aggregate rural efficiency units next period, and the (stochastic) transitory productivity shock and experience level. If the household is inexperienced, then she stays inexperienced. If the household is experienced, she stays that way with probability π and becomes inexperienced with probability $1 - \pi$.

The value function associated with a permanent move is:

$$v(a, r, s, x, i, N_r |\, \text{perm}) = \max_{a' \in \mathcal{A}} \left[u(Ra + w_r(z, s, i, N_r) - a' - m_p) + \beta\, \mathbb{E}[v(a', u, s', x', i', N_r')] \right] \tag{8}$$

While similar to the staying value function, there are several points of difference. First, the agent must pay m_p to make the permanent move and this costs resources. Second, continuation value functions denote the household's location changes from the rural area to the urban area. Third, with probability λ the household stays inexperienced, and with probability $1-\lambda$ the household becomes experienced.

The value function associated with a seasonal move is:

$$v(a, r, s, x, i, N_r |\, \text{seas}) = \max_{a' \in \mathcal{A}} \quad [u(Ra + w_r(s, i, N_r) - a' - m_T) + \beta\, \mathbb{E}[v(a', seas, s', x', i', N_r')]] \tag{9}$$

If a household decides to seasonally move, it pays the moving cost m_T, and it works in the urban area next period. The key distinction between the permanent move and the seasonal move is that this move is just for one period. The value function associated with a seasonal move is:

$$v(a', seas, s', x', i', N_r') = \max_{a'' \in \mathcal{A}} \left[u(Ra' + w_u(z, s') - a'')\bar{u}^{x'} + \beta\, \mathbb{E}[v(a'', r, s'', x'', i'', N_r'')] \right]. \tag{10}$$

There are several important points to take note of. First, this household only has one choice and that is how to adjust its asset holdings. By definition, the household engaging in a seasonal work in the urban area for only one period and then returns to the rural area. Second, this household experiences a disutility from the urban area as indicated by the presence of $\bar{u}$.

Examining (9) and (10) illustrates the forces that shape the seasonal move and, in turn, our inferences from the experimental and survey results. Generally, the choice to seasonally move is going to relate to a households comparative earnings advantage in the urban area relative to the rural area. However, there are several forces that may prevent this move. First, the urban disutility may prevent the household from moving even though its comparative advantage in the urban area is expected to be high. Second, there is risk associated with the move. A household does not know s' and hence there is a chance that the income realization in the urban area will not be favorable. Third, the household may have limited assets that simply make a move infeasible or not sufficient to insure against a bad outcome in the urban area.

Urban Households. Urban households face similar problems to those described above, though urban households choose between just staying or making a permanent move. For a household with productivity level z, the problem is:

$$v(a, u, s, x, N_r, i) = \max \left\{ v(a, u, s, x, N_r, i | \text{ stay}),\ v(a, u, s, x, N_r, i | \text{ perm}) \right\}. \tag{11}$$

Conditional on staying in the urban area, the value is:

$$v(a, u, s, x, i, N_r, | \text{ stay}) = \max_{a' \in \mathcal{A}} \left[u(Ra + w_u(z, s) - a') + \beta\, \mathbb{E}[v(a', u, s', x', i', N_r')] \right]. \tag{12}$$

Households staying in the urban area have three key differences from those staying in the rural area. First, their wage depends on their permanent productivity level, z, and not the season or number of aggregate efficiency units in the rural areas. Second, productivity shocks may have more or less volatility in terms of productivity, working through the γ parameter. Third, households that are inexperienced stay that way in the next period with probability λ and become experienced with probability $1 - \lambda$. This is one of the two outcomes over which expectations are taken in the Bellman equation; the other is the transitory shock, s.

The value function a permanent move back to the rural area is:

$$v(a, u, s, x, i, N_r | \text{ perm}) = \max_{a' \in \mathcal{A}} \left[u(Ra + w_u(z, s) - a' - m_p) + \beta\, \mathbb{E}[v(a', r, s', x', i', N_r')] \right]. \tag{13}$$

Here, the agent must pay m_p to make the permanent move and this costs resources. Furthermore, the continuation value function denote the household's location changes from the urban to the rural area. After a permanent move to the rural area, experienced households keep their experience with probability π, and lose it with probability $1 - \pi$.

2.3. Discussion: Determinants of Migration and Location Choice

The model allows for a rich set of determinants of migration, and location choice more generally. While in the following section we allow the data to discipline which are the most important determinants in practice in our empirical setting, it is worth discussing them informally here first.

One clear determinant of migration in the model is the permanent urban productivity level, z, which captures comparative advantage in the urban area. All else equal, agents with higher values of z will have stronger incentives to locate in the urban area. Seasonality is also a very clear determinant of migration in the model. Since the growing season has higher productivity than the lean season, rural households will be more likely to migrate (seasonally or permanently) to the urban area in the lean season, all else equal. The migration disutility, $\bar{u}$, is also an unambiguous deterrent to migration. The higher is $\bar{u}$, the less likely households will be to locate in the urban area. Finally, migration experience, x, is a positive determinant of migration, since experienced households face no disutility of locating in the urban area.

What role do the experience gain and loss terms, λ and π, play in migration and location decisions? Since all rural-urban migrants eventually become experienced, location decisions in the long run are not affected by λ and π. Instead, these terms mostly affect the extent of *repeat* migration. When experience is easy to obtain and hard to lose, i.e. λ is low and π is high, a subsidy to migration will induce inexperienced rural-urban migrants to repeat migrate (or stay in the urban area) for many periods in the future. For rural households induced to seasonally migrate, the lower is π, the less likely they will be to migrate in subsequent periods, since experience is lost at a faster rate.

The transitory shock, s, and asset levels, a, have ambiguous effects on migration and location choice. First suppose shocks are persistent, so that households with a high shock today are more likely to receive a high shock one period hence, and that $\gamma > 1$, so that shocks are more volatile in the urban area. In this case, rural households may be more likely to migrate to the urban area after receiving a good shock. The asset holdings also play a role here. High values of assets allow for insurance, which may mean that households only migrate in this case when their assets are sufficiently high. One concrete story that our model allows for in this case is that households with high productivity – either because of high z or high s shocks – are "misallocated" in the rural area, due to insufficient buffer stocks of savings. If this is the case, subsidizing migration may induce these high-productivity households to migrate and to realize large consumption gains due to a better allocation of their urban-specific productivity.

Suppose instead that shocks are persistent, but that $\gamma < 1$, so shocks are more volatile in the rural area than in the urban area. In this case, rural households may be more likely to migrate when they have bad shocks than good shocks. Since migration is costly both in monetary terms and non-monetary disutility, households may only migrate when they are sufficiently unproductive and when their assets are too low to insure themselves against their current low productivity. In this case, subsidizing migration may induce these low-productivity households to migrate, and to realize large consumption gains due to the higher average productivity level of the urban area.

Whether this case applies, or whether induced migrants tend to be high-productivity workers with low assets, is determined by the data. More generally, the welfare effects of subsidizing migration depend on the data used to discipline the model quantitatively. In particular, the welfare gains to a migrant induced to migrant depend on the size of the transfer to the migrant and monetary cost of migrating, the expected consumption gains from migrating, the consumption risk faced by the migrant, the non-monetary disutility of migrating, and how likely households are to gain and subsequently lose their experience. We turn to this in the next section.

3. Parameterizing the Model and Quantitative Analysis

In this section, we parameterize the model using two broad sources of evidence. The first is a set of cross-sectional moments that we calculate using the nationally representative Household Income and Expenditure Survey (HIES) of Bangladesh from 2010. The second is the controlled migration experiment of Bryan, Chowdhury, and Mobarak (2014), which we replicate within the model. We then discuss how we use this information to parameterize the model.

3.1. Directly Chosen Parameters

We begin by assigning some parameter values directly. We choose a time period to be half a year, to allow us to have seasonal migration and seasonal variation in rural productivity. We choose the discount factor, β, equal to 0.95 at the half-yearly frequency. We are currently experimenting with other parameter choices here. We choose the half-yearly return on storage, R, to be 0.95, to capture the half-yearly inflation rate in Bangladesh. Note that our low value of R helps reconcile the low saving rate we observe, much in the same way that Donovan (2016) reconciles low fertilizer inputs among rural agricultural workers with a negative observed return on storage of agricultural output. We choose the risk aversion parameter α to be 2.0, which is within the range of commonly chosen values in macro.

Seasonal variation in rural productivity is set so that the "lean" season is 65 percent less productive relative to the "growing" season. The seasonal moving cost is picked to be 10 percent of rural consumption. This approximately the cost of a round-trip bus ticket that the households in Bryan, Chowdhury, and Mobarak (2014) faced to seasonally migrate. We pick the permanently migration costs so that gross flows across regions are minor and not affecting our results. We set the degree of returns to scale in the rural area to be 0.1, which is consistent with the land share estimated by Valentinyi and Herrendorf (2008). Later we show that our model matches additional experimental variation on the extent to which wages rise when labor moves out of rural areas, from Akram, Chowdhury, and Mobarak (2017).

3.2. Productivity Processes

For the permanent productivity distributions, we assume that they are log normally distributed, and satisfying

$$z_r = 1, \text{ and } z_u \sim 1 - z^{-\theta}, \tag{14}$$

so that there is no variance in rural permanent productivity and the Pareto shape parameter θ controls the variance in urban productivity. The low variance of consumption in rural areas in the data pushes us towards having little or no variance in rural permanent productivity, and so we impose that directly here.

For the transitory shocks, we assume that they follow an AR(1) process:

$$\log s_{t+1} = \rho \log s_t + \epsilon_{t+1} \quad \text{with} \quad \epsilon_{t+1} \sim \mathcal{N}(0, \sigma_s).$$

3.3. Measurement Error in Income and Consumption Data

We allow for measurement error in income and consumption when we parameterize our model. In particular, we assume that rural consumption (that we observe directly using the data of Bryan, Chowdhury, and Mobarak (2014) satisfies:

$$\log \hat{c}_{r,i} = \log c_{r,i} + \log \upsilon_{r,i} \tag{15}$$

where $\hat{c}_{r,i}$ is observed consumption of individual i, $c_{r,i}$ is actual consumption, and $\upsilon_{r,i}$ is measurement error, which we assume is log normally distributed with variance $\sigma_{c,r}$.

Urban income (that we observe from the 2010 HIES) in turn satisfies:

$$\log \hat{y}_i = B_u + \log z_{u,i} + \log s_{u,i} + \log \upsilon_{u,i} \tag{16}$$

where $\hat{y}_i$ is actual income of i, B_u is a constant and $\upsilon_{u,i}$ is measurement error, which we assume is log normally distributed with variance $\sigma_{y,u}$.

The reason for these assumptions is that income and consumption in the data are clearly measured with error, and hence part of the income and consumption variance we observe arise from error, rather than permanent or temporary productivity variance. Thus, we allow for this possibility in the parameterization, treating $\sigma_{c,r}$ and $\sigma_{y,u}$ as unknowns.

3.4. Performing the Field Experiment in Model

The most novel feature of our calibration procedure is that we replicate the controlled migration experiment of Bryan, Chowdhury, and Mobarak (2014) directly in our model. In short, that experiment gave a cash transfer conditional on migration to individuals in a randomly selected subset of one hundred villages in the rural area of Rangpur in northern Bangladesh. In a set of control villages, around 36 percent of households sent a temporary migrant to work during the lean season. In the randomly assigned treatment villages, around 58 percent of households sent a temporary migrant. Of those in the treatment group that sent a migrant, consumption increased by 30 percent higher. Formally, this is estimated as a local average treatment effect (LATE) using instrumental variables, computed using assignment to the treatment group as an instrument for migration in the first stage, and regressing consumption on (instrumented) migration in the second stage. In the year after the experiment, migration rates were also higher in the treatment group, at 47 percent, compared to 36 percent in the control group. Thus, some of the individuals induced to migrate by the experiment returned of their own volition a year after the experiment, though many did not.

To replicate the experiment in the model, we solve for optimal policies of households who are faced with a one-time, unanticipated seasonal migration opportunity with out m_T. We randomly sample rural households in the model's stationary distribution. We then assign half to represent the control group, and half to represent the treatment group. We compute the fraction of model agents that migrate when given this conditional transfer of m_T, and compare that to the fraction that migrate in the control group. We then compute the LATE of migration on consumption as in the experiment, and the repeat-migration rate a year after the experiment.

3.5. Matching Moments in Model and Data

We parameterize the model using simulated method of moments. The nine moments we match are listed in Table 4 and can be divided up into three basic groups: experimental moments (4) and cross-sectional moments (5). The experimental moments are the fraction of migrants in the control group (36 percent), the fraction of migrants in the treatment group (58 percent), the consumption of those induced migrants relative to average consumption in the control (the LATE), and the repeat migration rate (9 percent). The cross-sectional moments are the urban-rural wage gap (1.80), the percent of households in the rural area (63 percent), the variance of log consumption in the rural area of non-migrants (0.12), the variance of log income in the urban area (0.68) and the percent of rural households with no liquid assets (47 percent).

We have nine moments that we use to target the moments above. These are the average productivity in the urban area, A_u, the shape parameter controlling permanent productivity differences θ, the standard deviation of transitory shocks, σ, the urban relative risk parameter γ, the autocorrelation of those shocks, ρ, the disutility of migration, $\bar{u}$, and the probability that a migration becomes experience (or becomes inexperienced), λ. Table 4 presents the values of the moments for the calibrated values of the parameters. In general, the model's predicted moments are quite similar to its counterparts in the data. Table 3 shows the estimated parameter values.[5]

There are a couple features of Table 3 worth pointing out. First, the shape parameter controlling permanent differences in ability is relative small, at around two. This implies that there is relatively large variation in permanent productivity in the urban area. Second, the urban "relative risk parameter" is around one, implying that a shocks in the urban area are roughly equally as volatile as in the rural area. Third, the disutility of the urban area is sizeable. Given the relatively large θ parameter, the large change in consumption (i.e. the LATE estimate) pushes to infer that the disutility associated with the urban area. We are mindful about simply inferring "residual wedges" from data and, hence, the next section uses new data we collected to asses to the plausibility of our model's predictions.

In order to get a better sense of how the model is identified, Table 6 presents the elasticities of each targeted moment to each parameter. To compute these elasticities, we follow Kaboski and Townsend (2011) and compute the percent change in each targeted moment for a percent change in each parameter, starting from the calibrated moments in Table 3. For expositional purposes we put in bold any elasticity greater than one in absolute value.

It is useful to discuss the results in Table 6 one moment at a time. The urban-rural wage gap is most responsive to θ, the productivity dispersion term, with all other parameters playing a modest role. This implies that the main determinant of the urban-rural wage gap, at least in the vicinity of the calibrated moments, is the extent to which high-productivity workers lo-

[5]Note that our method-of-moments procedure identifies the shock process to productivity to match all our moments, including the asset distribution and migration flows, which is similar to the approach used by Castañeda, Díaz-Gimenéz, and Ríos-Rull (2003). This is in contrast to approaches that e.g. estimate the shock process solely using observed wage variation in panel data. The fact that the observed shock process are obscured by the choice of location, is similar to the mechanisms of Banerjee and Newman (1993) and Munshi and Rosenzweig (2016).

cated in the urban area are more productive than rural workers. The rural employment share is most responsive to A_U and $\bar{u}$, the productivity in the urban area and disutility of moving there, which is as one would expect. The urban income variance is most responsive to θ, which is also intuitive, since higher productivity variation implies higher wage variation. The fraction of households with no assets is most responsive to σ, the variance of productivity shocks, which is because greater shocks implies that households do more self-insurance through saving.

The three migration rates – i.e. that of the control group, the treatment group, and the treatment group in year 2 – are affected in similar ways by the same basic set of parameters. Higher θ leads to lower migration rates, since there are fewer infra-marginal households that can expect to gain from migrating. Higher A_U means more migration, since the urban area features higher average wages, and higher $\bar{u}$ means lower migration, since the utility cost of migrating is higher. The migration rate in year 2 is also strongly affected by λ, which governs how quickly workers become accustomed to the urban area. Note that λ has the largest influence on the migration rate in year 2 of any other moment.

The effect of migration of consumption (i.e. the LATE reported in Bryan, Chowdhury, and Mobarak (2014)) is most responsive to $\bar{u}$, with all other parameters having little or no effect. This implies that the consumption gains from migration are mostly informing the size of the disutility of migration to the urban area. Intuitively, the migration gains from migrating in the model largely reflect the disutility of migrating. Finally, the rural consumption variance is most sensitive to σ and ρ, which is as one may expect.

Figure 1 plots the difference in migration rates between the treatment and controls groups in the model and data in 2008 (the year of the experiment in the model), and for five subsequent years. The model matches the experiment year and subsequent year as part of the calibration. As the figure shows, the model does well in other years as well, capturing the declining pattern present in the data. By five years after the experiment, the difference in migration rates between the two groups is positive but small in magnitude in the model, at 2 percent, and statistically insignificant in the data.

3.6. Policy Functions

In this section we discuss how the policy functions for location choice depend on the permanent productivity, z, as well as asset holdings, a, the transitory shock, s, and the experience level, x. We focus on the lean season, since most of the migration then, and for number of efficiency units in the rural area, N_r, in the stationary equilibrium.

Figure 2 plots the policy functions of a worker with a "moderate" level of urban productivity,

z. The x-axis represents the transitory productivity shock, and the y-axis is the asset holdings of the household. The upper panel plots the policy functions for a worker without migration experience. The dark blue region represents the set of a and s values such that the household permanently moves to the urban region; the medium blue region is where the household temporarily moves to the urban area; the light blue region is where the household stays in the rural area.

For sufficiently high asset holdings or transitory shocks, the household stays in the rural area. It is only the bottom half of the transitory shock distribution that leads to seasonal migration, and only then when assets are at a low level. The permanent moves come only for the lowest range of shocks and higher stocks of assets, for which the household can actually afford the permanent move. Thus, it is the workers who are the least selected on transitory shocks who migrate.

The bottom panel of Figure 2 shows the policy function for location choice of a household with the same z level but with experience. As one might expect, this leads to a much larger range of assets and transitory shocks for which the household makes either a temporary or permanent move. Experienced households with the lowest shock values always migrate, for any asset level. For higher shock values all but the households with the highest asset levels migrate. As for inexperienced households, those with relatively high shocks never migrate, regardless of the asset levels. Thus, even among households with experience, rural-urban migrants tend to be negatively selected on transitory shocks and asset holdings.

How do migration policies depend on permanent productivity, z? As one might expect, there is a clear positive relationship between z and rural-urban migration rates. As z increases, a much larger fraction of the s and a values correspond to a permanent or seasonal move. This has at least two important implications. First, the overall relationship between a worker's current productivity level (the product of its z and s) has an ambiguous relationship with migration rates. Higher z values increase the probability of migration, all else equal, while higher s values lower the probability of migration. Second, the set of workers in the rural area at any time period is likely to reflect the relatively low z values, since the high z types sort into the city. Thus, any policies directed towards workers in the rural area will be directed more towards households with low permanent productivity levels in the urban area. That is, the workers in the rural area at any given time are disproportionately those with a comparative disadvantage in the urban area.

3.7. How is the Model Identified?

In this section, we discuss how the experimental and cross sectional moments help identify the model. To do so, we take two alternative approaches. The first is to re-calibrate the model repeatedly each time restricting one parameter so as to shut off one particular channel, such as disutility of migration or differential risk between the two regions. The second approach is to start from the benchmark calibration, and then to compute the elasticity of each targeted moments to each parameter.

Table 5 presents the target moments in the data and in the benchmark calibration, as well as the best fit of the targeted moments under four alternative calibrations. These are (from column 4 to column 7) (i) when taking a limit as $\theta = \infty$, so that worker heterogeneity is shut down, (ii) when setting $\overline{u} = 1$, so that there is no disutility of migration, (iii) when setting $\lambda = 1$ so that workers never become experienced, and (iv) when $\gamma = 1$, so that there is no differential risk between the urban and rural areas. We discuss each calibration in turn.

When worker heterogeneity is shut down, the model does much worse on the baseline seasonal migration rate (48 percent versus 36 percent in the data) and experimental migration rate. The OLS coefficient of migration on consumption is way off (51 percent versus 10 percent in the data). Intuitively, this means that the "marginal households" are hard to get right without worker heterogeneity. The parameter θ helps the model get the right number of households near the margin. The migration rate in the control group, and the experimental migration rate, therefore are informative about the model's value of θ.

When migration disutility is shut down, the consumption gain for induced migrants is far too low (6 percent in the model versus 30 percent in the data). Similarly, the OLS coefficient of migration on consumption is too low (2 percent versus 10 percent in the data). This highlights how it is hard to match such large increases in consumption for induced migrants without disutility of migration in the model. Put differently, if there is no disutility of migration, it is hard to reconcile why workers require incentives to migrate and raise their consumption by 30 percent. In terms of identification, this means that the LATE of migration on consumption is informative about the model's value of $\overline{u}$.

When the effect of experience on migration disutility is shut down, the model's repeat migration rate is off (0 percent versus 9 percent in the data). Households in the model simply will not repeat migrate after being incentivized in the previous period, since the cost of migration is the same in the second period, but the incentives have been removed. The parameter λ is therefore necessary to get this repeat migration rate correct, and the repeat migration is the main moment in the data that informs the calibrated model's value of λ.

Finally, when differential risk is shut down, the OLS coefficient of migration on consumption is too high (20 percent versus 10 percent in the data). In the data, the experimental effect of migration on consumption is much higher than the OLS coefficient on migration from a regression of consumption on migration. This is consistent with the set of workers who migrate without incentives being negatively selected on income. When $\gamma = 1$, the OLS coefficient is too high, meaning that there is not enough negative selection in the model relative to the data. Thus, the model requires a lower γ coefficient to match the data. Put differently, the γ term is informative about the extent to which rural-urban migrants are positively or negatively selected on income relative to all workers in the rural area.

3.8. Welfare Predictions of the Model

We now discuss the model's predictions for the welfare implications of encouraging migration through conditional migration transfers. We being by presenting the welfare gains overall, and then go beneath the surface to look at welfare by asset and productivity levels. We compute welfare as the consumption-equivalent welfare metric used in macroeconomics since Lucas, Jr. (1985) and are the focus of recent studies such by e.g. Heathcote, Storesletten, and Violante (2010) and Krueger, Mitman, and Perri (2016). Concretely, the metric computes the percent increase in consumption, p, that makes the household indifferent between a p-percent consumption increase in perpetuity and being offered the conditional migration transfer.

Table 7 reports the overall average welfare gains by quintile of the income distribution. Income here takes into consideration both the permanent productivity and temporary shock, as well as the location of the household. As a frame of reference, we also report the welfare gains from an unconditional transfer system costing the same total amount. This amounts to giving a smaller amount of the consumption good to a larger number of households, without any conditions, and only to households in the rural area.[6]

Overall, the welfare gains are similar in magnitude for the conditional migration transfer and the unconditional transfer program. For the poorest quintile, however, the gains are about 25 percent larger for the migration transfers (1.42 percent versus 1.13 percent). As we show in more detail below, this is because the poorest households are substantially more likely to migrate in response to the conditional transfer. For households in the 2nd quartile, welfare gains are similar, at 0.52 percent and 0.58 percent consumption equivalent. For the 3rd, 4th and 5th quintiles, the welfare gains from either program are smaller in magnitude, and higher for the unconditional transfer. In the 5th quintile (i.e. the richest quintile), the conditional migration transfer induces very little migration, and leads to a negligible 0.07 percent welfare equivalent on average. The unconditional transfer affects many more households,

[6] There is a growing literature on the merits of conditions on transfers relative to unconditional transfers; see e.g. Baird, McIntosh, and Ozler (2016) and the references therein. In contrast to the setting of Baird, McIntosh, and Ozler (2016), where conditions make the extreme poor worse off, since they are too poor even to satisfy the conditions of the transfer, i.e. sending their children to school. In contrast, in our environment, even the poorest can migrate given the transfer, and moreover, have the strong incentive to migrate.

though by a small amount, and leads to a 0.19 percent welfare increase.

Figure 3 provides some more insight about how the welfare gains are determined, by plotting the consumption-equivalent welfare metric across all income and asset levels. The top panel of the figure plots the welfare gains (the colored surface) for the conditional migration transfer, while the bottom panel plots the welfare gains for the unconditional transfer program.

As the figure shows, it the households with the lowest income and asset levels that gain the most from the conditional migration transfers relative to the unconditional transfers. The key reason is that, as suggested in the policy functions of Figure 2, it is low asset holdings and low transitory shocks that lead to the most migration. When households are offered the conditional transfers to migrate, it is those with the lowest assets and worst shocks at present that have the highest take-up rates.

Why are the consumption-equivalent welfare gains from migration not higher, if the workers induced to migrate gain 30 percent consumption on average? There are several forces at work in the model that determine the model's welfare impacts. First, as we showed in Table 5, the model requires two features to match the consumption gains for induced migrants and the extent of persistence in repeat migration from the experiment. The first is a relatively high non-monetary disutility of migration for inexperienced migrants, coming through the $\overline{u}$ term. The second is a relatively high probability of losing migration experience. Putting these together, the model implies that the monetary gains from consumption are in large part offset by non-monetary disutility of migration that is not easy to overcome for those returning to the rural area each year.

Relatedly, the model shows that those induced to migrate are those that are negatively selected on both permanent urban productivity and temporary productivity shocks. Thus, the model tells us that it is not predominantly high-productivity workers who are "misallocated" in the rural area and unable to migrate, due to say insufficient assets to pay for the moving cost, or insufficient assets to insure against the greater urban risk. These workers would not require a high disutility of migration to keep them from migrating. Instead, the model and experimental data imply that there are not many of these high-z types stuck in the rural area. Therefore, the model requires that the induced migrants be the low-productivity households that pass up large consumption increases from migration, without incentives to migrate, due to non-monetary costs of migrating.

3.9. Discussion: Misallocation across Rural and Urban Areas

The large and growing literature on misallocation and TFP has found evidence of large potential misallocation in developing countries, such as in Indian and Chinese manufacturing (Hsieh and Klenow, 2009), and African and Chinese farming (Restuccia and Santaeulalia-Llopis, 2016; Adamopoulos, Brandt, Leight, and Restuccia, 2016). One large potential source of misallocation is workers across the agricultural and non-agricultural sectors Gollin, Lagakos, and Waugh (2014) or alternatively between rural and urban areas (Young, 2013). While these papers don't discuss welfare per se, the idea is that there could be large welfare gains from reallocating factors of production across firms, or workers across space.

Our paper's findings suggest that, at least in Bangladesh, the gains from encouraging workers out of rural areas are likely to be modest. Urban areas do pay higher wages, but workers seem to face large disutility of migrating to urban areas, and those with the strongest comparative advantage in urban areas seem already to be there. In this sense, our interpretation of the apparent misallocation is closer to that of Lagakos and Waugh (2013), Young (2013) and Herrendorf and Schoellman (2016), who posit that existing allocations are close to efficient. The difference is that those studies focus exclusively on selection, while the current paper finds a role for selection and a large disutility of workers moving from rural to urban areas. In the section we turn to just what this disutility may represent.

4. Validating the Model with New Survey Evidence

One key prediction of the model is that the non-monetary disutility of migrating is large. This is important for our welfare predictions, because it implies that even when consumption of induced migrants rises by a substantial 30 percent, these consumption gains are offset in large part by disutility they experience from migrating. In this section, we assess the plausibility of this large non-monetary disutility of migration. To do so we conduct a new survey on migration on the same experimental sample of households used to discipline the model.

4.1. Discrete Choice Experiments using Hypothetical Migration Choices

To better understand the non-monetary costs of migration, we conducted a new discrete-choice experiment on the same sample of households used to calibrate our model. Since the field experiment of Bryan, Chowdhury, and Mobarak (2014) does not create random variation in the non-monetary costs of moving, we focused this new experiment around variation in non-monetary costs of migration, in order to estimate their value in monetary terms. We focus on the same sample of treatment and control households in the field experiment, The choice experiments presented a series of hypothetical scenarios to respondents in which we randomly varied a few key attributes associated with one of two migration options. The surveys asked respondents to indicate what migration choice they would make, when presented with (hypothetical) options for the fall 2015 lean season. The attributes we presented to respondents under each option randomly varied the probability of finding employment in the city, the wage if employed, how frequently the migrant could return to visit family (to minimize separation), and access to a hygienic latrine in their residence at the migration destination, which is a useful proxy for the quality of housing amenities the migrant would experience in the city. We conducted choice experiments because it would be impossible to vary all these attributes in a controlled manner in a field setting. Such discrete choice experiments (DCEs) are frequently used in marketing to estimate the effect of specific attributes on the attractiveness of a product to consumers, in environmental economics (and in environmental policy-making) to infer the value of environmental goods and services for which market transaction data do not exist (Hanemann, 1994), and in health economics to understand service provider and patient preferences.

Figure 4 presents one example of the choices we presented to the respondents. The respondent is asked to choose one of the two migration options presented, or a third ”opt-out” no-migration option.[7] The experimental setup for the hypothetical options should mimic the circumstances under which the equivalent decision would be made in the real world (Ryan and Skatun, 2004). In this example, both options feature a 33 percent chance of employment. Choice #1 offers a lower wage if employed but better amenities (more regular family contact and a hygienic latrine in the residence) compared to Choice #2.

We conducted these DCEs on a sample of 2,714 respondents, and each respondent is presented with 7 different choice sets for which the values of attributes are varied. We use the Choice Experiment tools in JMP12 (built on SAS) to generate algorithms that picks values for the attributes under each migration option in each choice problem is such a way that the power of the experiment is maximized. We observe a total of 18,998 choices, but to eliminate any bias stemming from recent induced migration experience, we only use choices made by respondents who reside in the control villages of the Bryan, Chowdhury, and Mobarak (2014) migration subsidy experiment. We estimate a multinomial logit model of migration choice as a function of the offered attributes of each location using these 3,349 observations. Table 8 presents the predicted probabilities and estimated marginal effects from this multinomial logit regression. We report the marginal effects of improving each attribute associated with option #2 on the probability of choosing option #2, setting all attributes associated with

[7]The methodological literature on DCEs strongly recommends that an opt-out option consistent with the decision at hand is always provided (Lancsar and Louviere, 2008).

option #1 at the least attractive values, and setting attributes associated with option #2 at median values. The rationale for evaluating the marginal effects on choosing destination #2 at the median values for 2 and minimum values for 1 is to effectively create only two relevant choices for the potential migrant: either migrate to destination 2, or stay at home. This binary choice most closely corresponds to the decisions made by agents in our theory, where we only model a binary migration choice.

The middle two data columns of Table 8 show the predicted probabilities (PP) and marginal effects (ME) on the propensity to migrate to destination #2 when the characteristics of destination #2 are varied. The first and last two data columns show the PP and ME on destination #1 and No Migration when the characteristics of Opportunity #2 are varied. Of the four attributes for each destination that we specified in our surveys, the probability of employment had three possible values: 33%, 66% and 100%, the daily wage had five possible values: 200, 235, 270, 305 and 340 taka per day, living conditions had two categories: *pucca* (hygienic) latrine in residence, or no latrine, and the disutility associated with separation from family had three possible categories: an ability to go back and visit family once, twice or 4 times during the two month seasonal migration period. The daily wage is modeled as a continuous variable in the multinomial logit, while the other attributes are modeled as categorical variables.

Table 8 shows that an increase in employment probability at destination 2 from 33% to 66% or 100% (holding destination #1 characteristics fixed) increases the propensity to migrate to destination #2 by by 13.5 and 19.3 percentage points respectively. Unemployment risk is therefore a quantitatively important deterrent to migration. The next three rows show that the frequency of family visits has a negligible effect on migration choices. Having a latrine in residence increases the probability of choosing destination #2 by 17.4 percentage points. Housing conditions at the destination are an important determinant of migration choices. The probability of migrating to destination #2 increases by 0.4 percentage points for every additional Taka in daily wage that is on offer. In other words, the migration probability jumps by 20 percentage points if the destination offers an extra 50 Taka in daily income. Thus, having a better housing option is similar to an additional 50 Taka per day in wages, which is equivalent to around 20 percent of average wages at the origin. To the extent that rural-urban migrants generally face poor urban housing options, in particular those without latrines, this represents a large non-monetary cost of migration and a substantial offsetting force to the higher wages earned by migrants.

5. Conclusions and Future Work

This paper studies the welfare implications of subsidizing rural-urban migration in low-income countries. Cross sectional data show that most low-income countries have much higher wages in urban areas than in rural areas. One interpretation of these wage gaps is that workers are misallocated across space due to, say, credit constraints. If this is the case, one may expect substantial welfare gains from providing transfers to rural workers conditional on migrating to urban areas.

To estimate the welfare effects of encouraging migration, we build a dynamic model of migration that captures several main existing determinants of migration emphasized in the literature. We discipline the model using cross-sectional survey data, and a field experiment on migration from Bangladesh. The novelty in our parameterization is that we replicate the experiment within the model, and choose the model's parameters so as to match a host of experimental moments. In particular, our model matches the experimental elasticity of migration to a migration subsidy, the experimental consumption gains for induced migrants, and the degree of repeat migration in subsequent years for repeat migrants.

The parameterized model predicts that the average welfare gains from encouraging migration are modest in magnitude, and comparable to the average gains from unconditional transfers costing the same total amount. The reason the gains are not higher is that induced migrants tend to be negatively selected on income, and the model's non-monetary disutility of migration is large. We validate the former directly by comparing characteristic of all migrants to induced migrants, and the latter using a new discrete choice experiment that we conducted on the same experimental sample. Still, for the poorest households, the welfare gains from migration subsidies are higher than unconditional cash transfers costing the same total amount. This suggests that conditional migration transfers may be a useful way to raise the welfare of poor rural households in the developing world. Our study does not, however, point to a low-cost path to large welfare gains from better allocating workers across rural and urban areas, at least in Bangladesh. Our discrete choice experiments point to low-quality housing options for migrants in urban areas as a key non-monetary cost of migration. This suggests that investments in urban infrastructure may be an important input to large welfare gains from migration. Future research should explore the consequences of encouraging migration over space in other countries and settings, and the interactions between urban infrastructure and the welfare gains from migration.

References

ADAMOPOULOS, T., L. BRANDT, J. LEIGHT, AND D. RESTUCCIA (2016): "Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China," Unpublished Working Paper, University of Toronto.

ADAMOPOULOS, T., AND D. RESTUCCIA (2014): "The Size Distribution of Farms and International Productivity Differences," *American Economic Review*, 104(6), 1667–1697.

AIYAGARI, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109(3), 659–84.

AKCIGIT, U., H. ALP, AND M. PETERS (2016): "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries," Unpublished Working Paper, University of Chicago.

AKRAM, A. A., S. CHOWDHURY, AND A. M. MOBARAK (2017): "General Equilibrium Effects of Emigration on Rural Labor Markets," Unpublished Working Paper, Yale University.

ASKER, J., A. COLLARD-WEXLER, AND J. DE LOECKER (2015): "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 105(1), 131–171.

BAIRD, S., C. MCINTOSH, AND B. OZLER (2016): "Cash or Condition? Evidence from a Cash Transfer Experiment," *Quarterly Journal of Economics*, 126(4), 1709–1753.

BANDIERA, O., R. BURGESS, N. DAS, S. GULESCI, I. RASUL, AND M. SULAIMAN (Forthcoming): "Labor Markets and Poverty in Village Economics," *Quarterly Journal of Economics*.

BANERJEE, A. V., AND A. F. NEWMAN (1993): "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101(2), 274–98.

BAZZI, S., A. GADUH, A. ROTHENBERG, AND M. WONG (2016): "Skill Transferability, Migration, and Development: Evidence from Population Resettlement in Indonesia," *American Economic Review*, 106(9), 2658–98.

BEEGLE, K., J. DE WEERDT, AND S. DERCON (2011): "Migration and Economic Mobility in Tanzania," *Review of Economics and Statistics*, 93, 1010–1033.

BEEGLE, K., E. GALASSO, AND J. GOLDBERG (2016): "Direct and Indirect Effects of Malawi's Public Works Program on Food Security," Unpublished Working Paper, University of Maryland.

BEWLEY, T. (1977): "The Permanent Income Hypothesis: A Theoretical Formulation," *Journal of Economic Theory*, 16(2), 252–292.

BRUECKNER, J. K., AND S. V. LALL (2015): "Cities in Developing Countries: Fueled by Rural-Urban Migration, Lacking in Tenure Security, and Short of Affordable Housing," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. Strange, vol. 5B, pp. 1399–1451.

BRYAN, G., S. CHOWDHURY, AND A. M. MOBARAK (2014): "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh," *Econometrica*, 82(5), 1671–1748.

BRYAN, G., AND M. MORTEN (2015): "Economic Development and the Spatial Allocation of Labor: Evidence From Indonesia," Unpublished Working Paper, Stanford University.

BUERA, F. J., J. P. KABOSKI, AND Y. SHIN (2014): "Macro-Perspective on Asset Grants Programs: Occupational and Wealth Mobility," *American Economic Review Papers and Proceedings*, 104(5), 159–164.

BUERA, F. J., AND Y. SHIN (2013): "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 121(2), 221–272.

CALIENDO, L., M. DVORKIN, AND F. PARRO (2015): "Trade and Labor Market Dynamics," Unpublished Working Paper, Yale University.

CASELLI, F. (2005): "Accounting for Cross-Country Income Differences," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf.

CASTAÑEDA, A., J. DÍAZ-GIMENÉZ, AND J.-V. RÍOS-RULL (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111(4), 818–57.

DAVID, J., H. HOPENHAYN, AND V. VENKATESWARAN (2016): "Information, Misallocation and Aggregate Productivity," *Quarterly Journal of Economics*, 131, 943–1005.

DONOVAN, K. (2016): "Agricultural Risk, Intermediate Inputs, and Cross-Country Productivity Differences," Unpublished Working Paper, University of Notre Dame.

FAJGELBAUM, P. D., E. MORALES, J. C. S. SERRATO, AND O. ZIDAR (2015): "State Taxes and Spatial Misallocation," Unpublished Working Paper, UCLA.

GOLLIN, D., D. LAGAKOS, AND M. E. WAUGH (2014): "The Agricultural Productivity Gap," *Quarterly Journal of Economics*, 129(2), 939–993.

Greenwood, J., P. Kircher, C. Santos, and M. Tertilt (2013): "An Equilibrium Model of the African HIV/AIDS Epidemic," NBER Working Paper 18953.

Guner, Nezih, G. V., and Y. D. Xu (2008): "Macroeconomic Implications of Size-Dependent Policies," *Review of Economic Dynamics*, 11(4), 721–744.

Hall, R. E., and C. I. Jones (1999): "Why Do Some Countries Produce So Much More Output per Worker than Others?," *Quarterly Journal of Economics*, 114(1), 83–116.

Hanemann, M. W. (1994): "Valuing the Environment through Contingent Valuation," *Journal of Economic Perspectives*, 8(4), 19–43.

Harris, J. R., and M. P. Todaro (1970): "Migration, Unemployment and Development: A Two-Sector Analysis," *American Economic Review*, 60(1).

Heathcote, J., K. Storesletten, and G. L. Violante (2009): "Quantitative Macroeconomics with Heterogenous Households," *Annual Reviews in Economics*, 1(1), 319–354.

——— (2010): "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 118(4), 681–722.

Herrendorf, B., and T. Schoellman (2016): "Wages, Human Capital, and Structural Transformation," Unpublished Manuscript, Arizona State University.

Hnatkovska, V., and A. Lahiri (2013): "Structural Transformation and the Rural-Urban Divide," Unpublished Working Paper, University of British Columbia.

Hsieh, C.-T., and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403–1448.

Huggett, M. (1996): "Wealth Distribution in Life-Cycle Economies," *Journal of Monetary Economics*, 38(3), 469–94.

Kaboski, J. P., and R. M. Townsend (2011): "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative," *Econometrica*, 79(5), 1357–1406.

Kaplan, G., and S. Schulhofer-Wohl (2017): "Understanding the Long-Run Decline in Interstate Migration," *International Economic Review*, 58(57-94), 211–251.

Kaplan, G., and G. L. Violante (2010): "How Much Consumption Insurance Beyond Self-Insurance?," *American Economic Journal: Macroeconmoics*, 2(4), 53–87.

Kennan, J., and J. R. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79(1), 211–251.

KHANDKER, S. R. (2012): "Seasonality of Income and Poverty in Bangladesh," *Journal of Development Economics*, 97(2), 244–256.

KRUEGER, D., K. MITMAN, AND F. PERRI (2016): "On the Distribution of the Welfare Losses of Large Recessions," Unpublished Working Paper, University of Pennsylvania.

LAGAKOS, D., AND M. E. WAUGH (2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *The American Economic Review*, 103(2), 948–980.

LANCSAR, E., AND J. LOUVIERE (2008): "Conducting Discrete Choice Experiments to Inform Healthcare Decision Making," *PharmacoEconomics*, 26(8), 661–677.

LUCAS, JR., R. E. (1985): "Models of Business Cycles," in *Lectures on Economic Growth*. Oxford: Basil Blackwell.

MCKENZIE, D., J. GIBSON, AND S. STILLMAN (2010): "How Important is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration," *Journal of the European Economic Association*, 8(4), 913–945.

MCMILLAN, M. S., AND D. RODRIK (2011): "Globalization, Structural Change and Productivity Growth," NBER Working Paper No. 17143.

MIDRIGAN, V., AND D. XU (2014): "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104(2), 422–458.

MIGUEL, E., AND J. HAMORY (2009): "Individual Ability and Selection into Migration in Kenya," Unpublished Manuscript, University of California, Berkeley.

MOLL, B. (2014): "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, 104(10), 3186–3221.

MORTEN, M. (2013): "Temporary Migration and Endogenous Risk Sharing in Village India," Unpublished Working Paper, Stanford University.

MUNSHI, K., AND M. ROSENZWEIG (2016): "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *American Economic Review*, 106(1), 46–98.

MURALIDHARAN, K., P. NIEHAUS, AND S. SUKHTANKAR (2016): "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India," Unpublished Working Paper, University of California San Diego.

PETERS, M. (2016): "Heterogeneous Mark-Ups, Growth and Endogenous Misallocation," Unpublished Working Paper, Yale University.

RAUCH, J. E. (1993): "Economic Development, Urban Underemployment, and Income Inequality," *Canadian Journal of Economics*, 26(4).

RESTUCCIA, D., AND R. ROGERSON (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11(4), 707–720.

RESTUCCIA, D., AND R. SANTAEULALIA-LLOPIS (2016): "Land Misallocation and Productivity," Unpublished Working Paper, University of Toronto.

RESTUCCIA, D., D. T. YANG, AND X. ZHU (2008): "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis," *Journal of Monetary Economics*, 55, 234–250.

ROY, A. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–46.

RYAN, M., AND D. SKATUN (2004): "Modelling Non-demanders in Choice Experiments," *Health Economics*, 13(4), 397–402.

TODD, P. E., AND K. I. WOLPIN (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.

VALENTINYI, A., AND B. HERRENDORF (2008): "Measuring Factor Income Shares at the Sectoral Level," *Review of Economic Dynamics*, 11, 820–835.

VOLLRATH, D. (2009): "How Important are Dual Economy Effects for Aggregate Productivity?," *Journal of Development Economics*, 88(2), 325–334.

YANG, M.-J. (2016): "Micro-level Misallocation and Entry Selection," Unpublished Working Paper, University of Washington.

YOUNG, A. (2013): "Inequality, the Urban-Rural Gap and Migration," *The Quarterly Journal of Economics*, 129(2), 939–993.

Tables and Figures

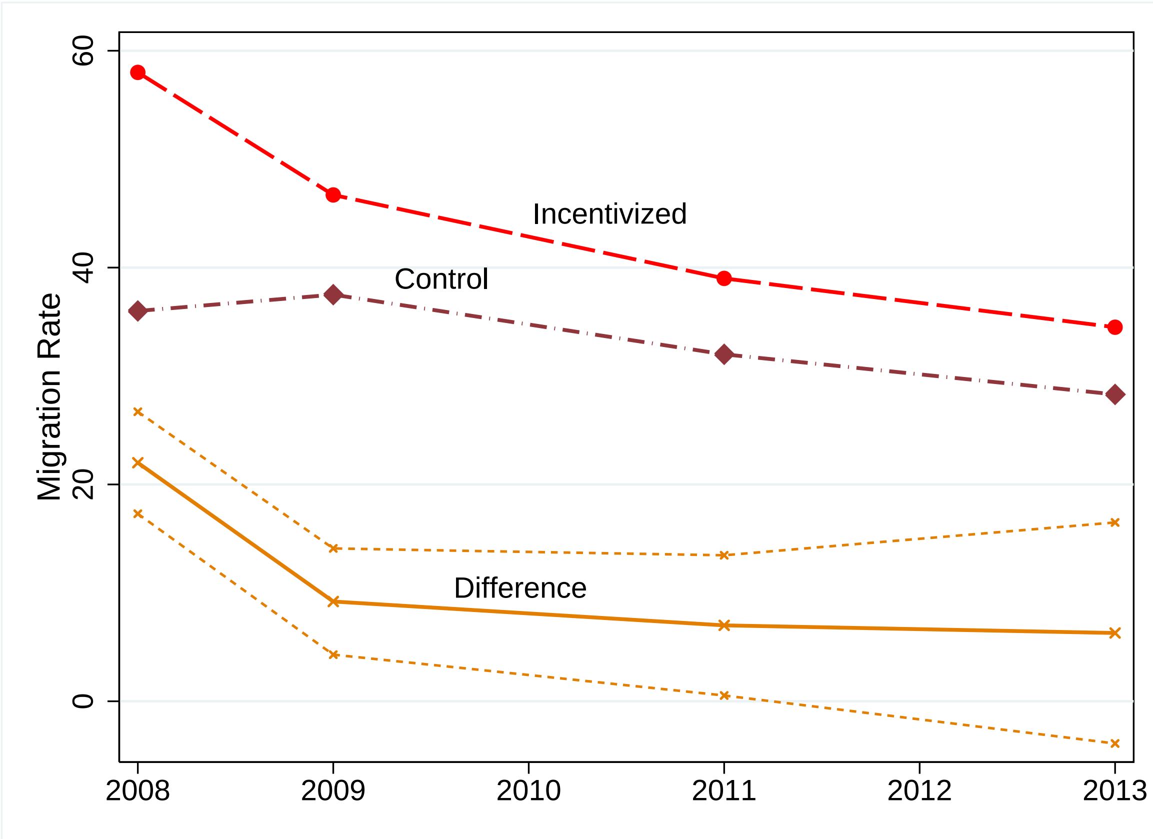
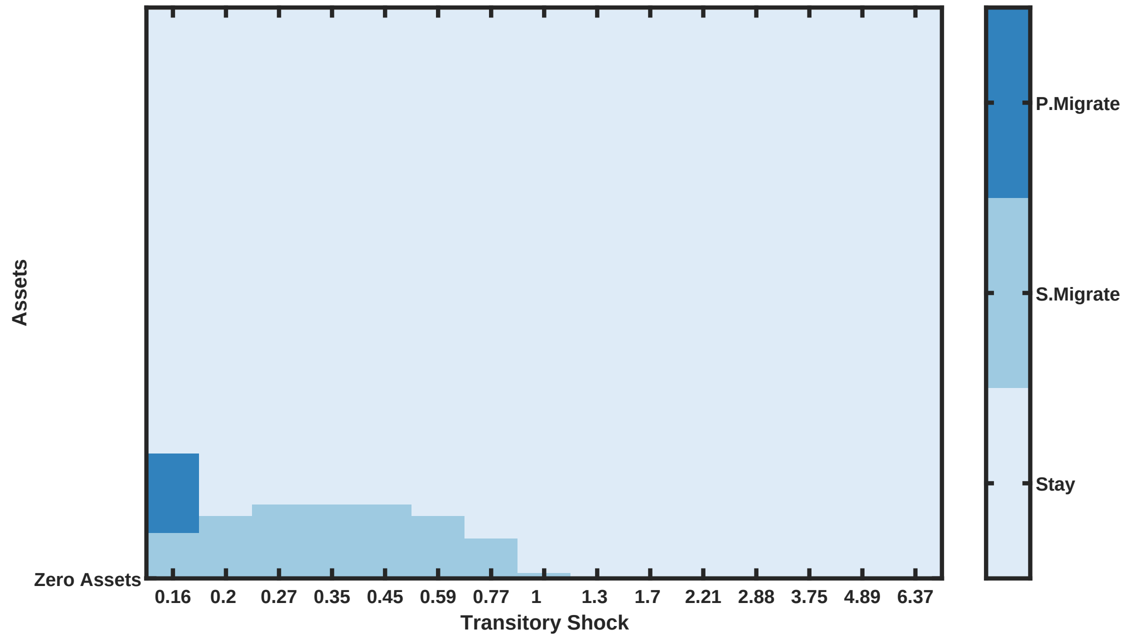
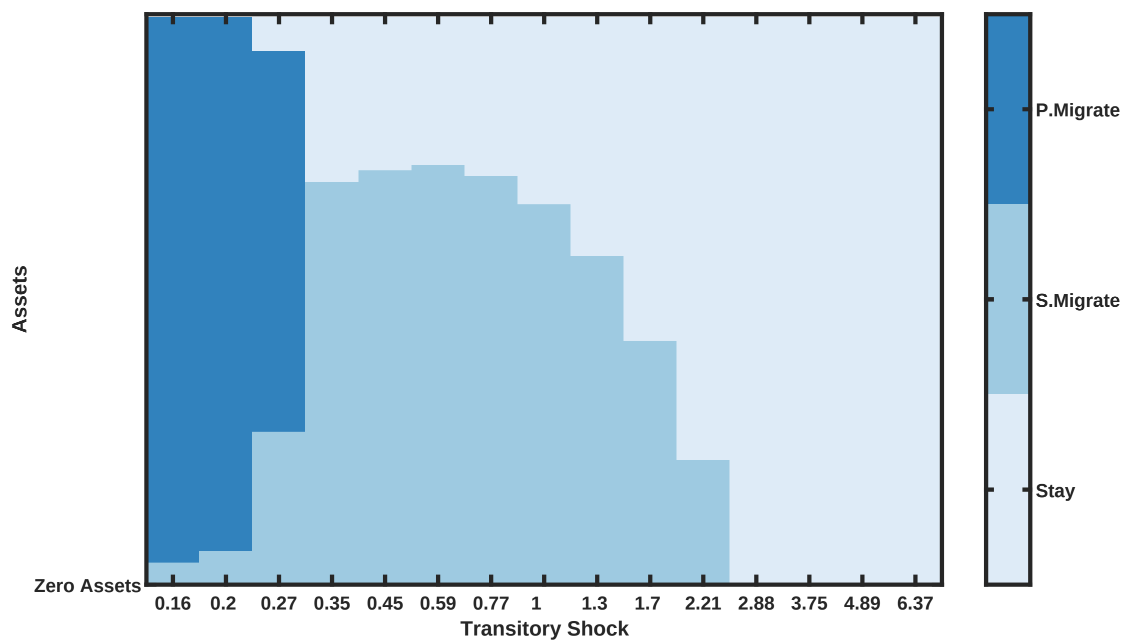


Figure 1: Migration Rates in Treatment and Control Groups

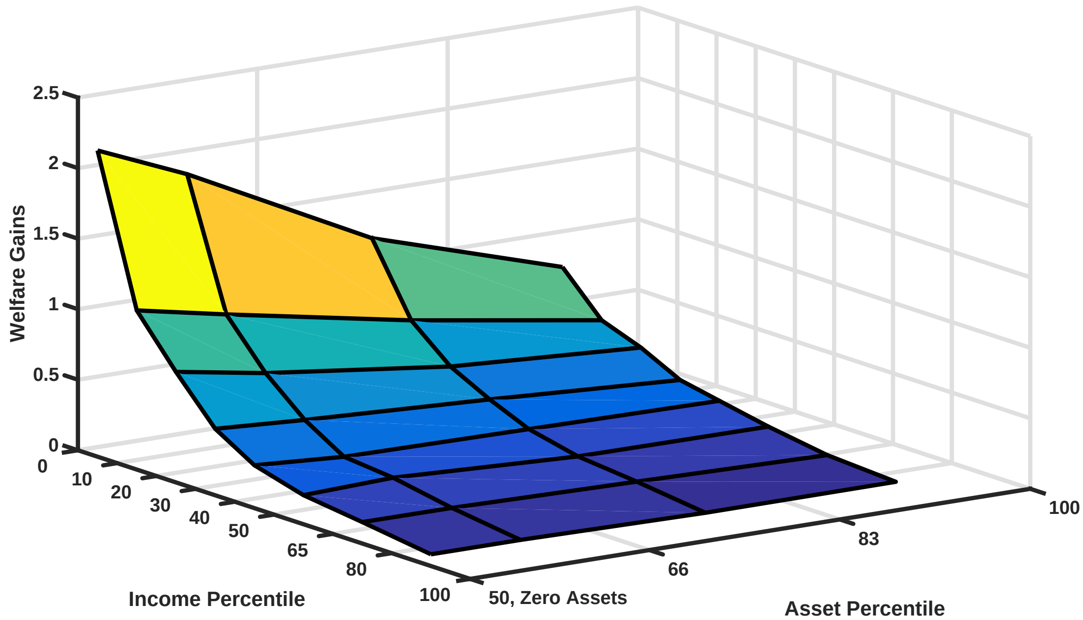


(a) Lean Season, No Experience

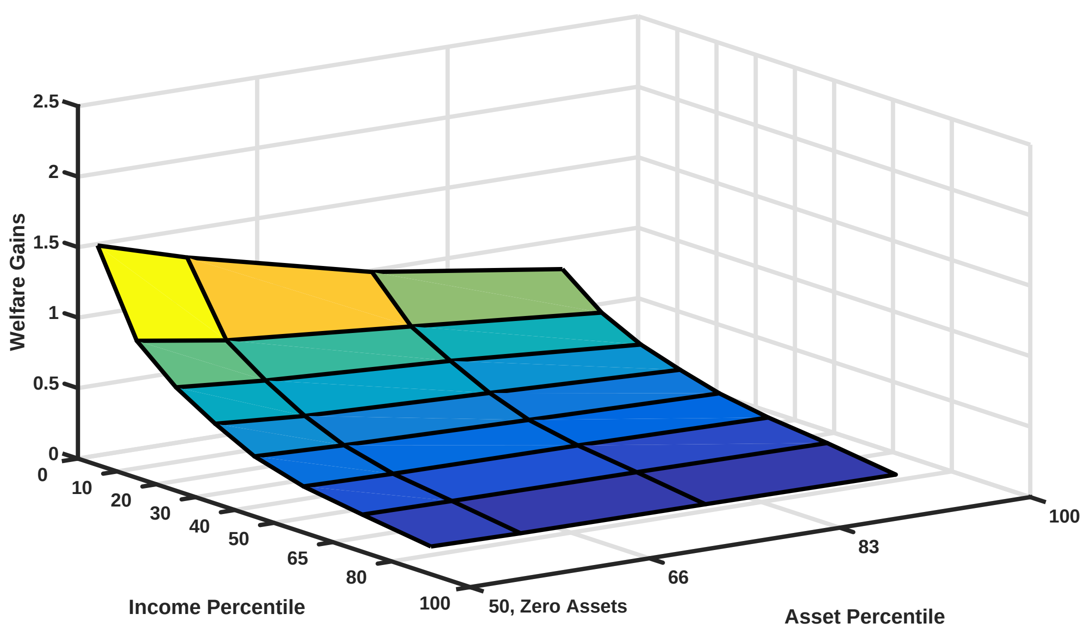


(b) Lean Season, With Experience

Figure 2: Migration Policies for Household with Intermediate z



(a) Migration Subsidy



(b) Unconditional Transfer

Figure 3: Consumption-Equivalent Welfare Gains

S.1.C.2			
Given the attributes below, which option do you choose? Please evaluate each new pair of migration options independent of the ones you saw earlier.			
	Choice #1: Migration	**Choice #2: Migration**	**Choice #3: No Migration**
Chance of Employment	33%	33%	N/A
Daily Wage (Taka)	270	340	Wage at Home in November
Latrine Facility during Migration	Pucca Latrine in Residence	Walk to Open Defecate or Public Pay Toilet	N/A
Family Contact	See Family Every Month	See Family Every 2 Month	N/A
s16bq2_1 Your Choice *(Tick Single Box)*	☐	☐	☐

Figure 4: Sample Migration Opportunity

Table 1: Urban-Rural Wage Gaps in Bangladesh

	Mean	Median
All urban / All rural	1.80	1.72
All urban / Rangpur rural	2.12	1.87
Dhaka / Rangpur rural	2.20	2.01

Note: The table reports the ratio of average wages in urban to rural areas in Bangladesh in 2010 with alternate subsets of rural and urban areas. In all rows, wages are calculated as monthly labor earnings divided by hours worked. The sample is restricted to workers with positive wage earnings that are at least 15 years old and who worked at least six months in the last year. The wages are computed using the 2010 HIES of Bangladesh.

Table 2: Parameters Chosen Directly

Parameter	Value	Source
Time period	Half year	—
Risk aversion, α	2.0	—
Discount factor, β	0.95	—
Gross real interest rate, R	0.96	Semi-annual inflation rate
Rural seasonal productivity ratio, A_{rl}/A_{rg}	0.65	Khandker (2012)
Seasonal moving costs, m_T	10 % of rural cons.	BCM (2014)
Permanent moving costs, m_p	$6 \times m_T$	1% Permanent Moves

Note: The table lists the parameters whose values were chosen directly, their values in the benchmark quantitative analysis and the rationale for their choice.

Table 3: Calibrated Parameters

Parameter	Value
Migration disutility, $\bar{u}$	1.49
Probability gaining experience, $1-\lambda$	0.44
Probability losing experience, $1-\pi$	0.56
Shape parameter, urban talent, θ	1.98
Urban relative shock, γ	0.60
Productivity urban area A_u	1.42
Standard deviation of transitory shocks	0.34
Persistence of transitory shocks	0.75

Note: The table lists the parameters whose values were determined using simulated method of moments, as described in the text, and their values in the quantitative analysis.

Table 4: Calibration Results: Model Fit

Moments	Data	Model
Control: Seasonal migrants	36	32
Control: Consumption increase of migrants (OLS)	10	10
Experiment: Seasonally migration relative to control	22	20
Experiment: Percent of treatment group that seasonally migrate in year 2	9	6
Experiment: Consumption of induced migrants relative to control (LATE)	30	30
Urban-Rural wage gap	1.80	1.80
Percent in rural	63	63
Variance of log wages in urban	0.68	0.68
Variance of log consumption of rural non-migrants	0.12	0.12
Percent of rural households with no liquid assets	47	47

Note: The table reports the moments targeted using simulated method of moments, and their values in the data and in the model.

Table 5: Calibration Results under Alternative Parameter Restrictions

Moments	Data	Baseline	$\theta = \infty$	$\bar{u} = 1$	$\lambda = 1$	$\gamma = 1$
Control						
Migration	36	32	48	29	37	26
Repeat Migration, year 2	25	25	31	27	25	25
Repeat Migration, year 4	16	18	16	25	18	23
Consumption gain, migrants (OLS)	10	10	51	2	10	20
Experiment						
Migration	22	20	31	10	16	18
Migration, year 2	9	6	0	0	0	06
Consumption gain, induced migrants (LATE)	30	30	20	6	30	30
Aggregate						
Urban-Rural wage gap	1.80	1.80	1.80	1.80	1.80	1.80
Percent in rural	63	63	57	89	92	58

Note: The table reports the results of the calibration in the benchmark case and in several alternative parameter restrictions.

Table 6: Elasticities of Targeted Moments to Parameters

	σ	θ	A_U	ρ	$\bar{u}$	λ	γ
Urban-rural wage gap	0.03	**-1.37**	-0.27	-0.03	0.00	-0.05	0.19
Rural employment share	0.00	-0.61	**-1.72**	0.00	**-1.13**	0.00	0.26
Urban income variance	0.90	**-2.10**	0.12	-0.35	0.08	0.00	0.88
Fraction of households with no assets	**-1.55**	-0.08	-0.62	0.39	0.08	-0.05	-0.07
Migration rate (control)	-0.57	**4.09**	**3.01**	0.58	**-10.75**	0.81	**-1.27**
Migration rate (treatment)	0.83	**3.68**	**4.31**	-0.87	**-3.04**	0.000	0.000
Migration rate (treatment, year 2)	0.84	**4.39**	**4.13**	**-1.31**	**-2.81**	**-1.24**	-0.39
Migration effect on consumption (LATE)	0.06	-0.23	-0.08	-0.04	**1.97**	-0.03	-0.19
Rural consumption variance	**2.10**	0.02	0.74	**-1.62**	-0.04	-0.02	-0.22

Note: This table reports the elasticities of each targeted moment to each parameter. Elasticities are calculated by computing the percent increase in each moment to a one percent increase in each parameter, starting from the calibrated parameters of the model. For expositional purposes, elasticities greater than one in absolute value are printed in bold.

Table 7: Consumption-Equivalent Welfare Gains by Income Quintile

	Income Quintile				
	1st	2nd	3rd	4th	5th
Conditional Migration Transfers	1.42	0.52	0.27	0.16	0.07
Unconditional Transfers (same total cost)	1.13	0.58	0.39	0.28	0.19

Note: The table reports the consumption-equivalent welfare gains from the conditional migration transfers relative to an unconditional transfer program costing the same total amount.

Table 8: Estimated Marginal Effects on Migration

	Migration Opp. #1		Migration Opp. #2		No Migration	
	PP	ME	PP	ME	PP	ME
33% Prob. Employment	0.116***	0.000	0.597***	0.000	0.286***	0.000
	(0.018)	(.)	(0.053)	(.)	(0.055)	(.)
66% Prob. Employment	0.067***	-0.049***	0.732***	0.135***	0.200***	-0.086***
	(0.012)	(0.010)	(0.046)	(0.030)	(0.045)	(0.029)
100% Prob. Employment	0.048***	-0.068***	0.791***	0.193***	0.161***	-0.125***
	(0.009)	(0.012)	(0.040)	(0.033)	(0.038)	(0.033)
Family visit once in 60 days	0.071***	0.000	0.760***	0.000	0.169***	0.000
	(0.014)	(.)	(0.041)	(.)	(0.040)	(.)
Family visit twice in 60 days	0.067***	-0.004	0.732***	-0.027	0.200***	0.032
	(0.012)	(0.008)	(0.046)	(0.024)	(0.045)	(0.023)
Family visit 4 times in 60 days	0.058***	-0.013*	0.763***	0.003	0.179***	0.010
	(0.012)	(0.007)	(0.049)	(0.028)	(0.046)	(0.028)
No Latrine in residence	0.067***	0.000	0.732***	0.000	0.200***	0.000
	(0.012)	(.)	(0.046)	(.)	(0.045)	(.)
Pucca Latrine in residence	0.029***	-0.038***	0.906***	0.174***	0.065***	-0.136***
	(0.006)	(0.008)	(0.021)	(0.032)	(0.019)	(0.032)
Daily Wage (Taka), Opp #2		-0.001***		0.004***		-0.002***
		(0.000)		(0.000)		(0.000)
Observations	3449	3449	3449	3449	3449	3449

Note: Standard errors are adjusted for 2,566 clusters in hhid. PP columns represent predicted probabilities of migrating at given condition, and ME columns represent marginal effects of chaning migration conditions in each category. PP and ME are measured while fixing 1st migration conditions (wage, employment chance, family visit, latrine) at the worst, and fixing 2nd migration condition at median. Analysis sample includes only those households in the control group.