Measuring Poverty and Vulnerability in Real-Time

Joshua Blumenstock*  Michael Callen†  Tarek Ghani‡
U.C. Berkeley  U.C. San Diego, NBER  Washington University

Niall Keleher§  Jacob Shapiro¶
U.C. Berkeley  Princeton University

December 21, 2016
PRELIMINARY RESULTS - NOT FOR PUBLIC CIRCULATION OR CITATION

Abstract
In wealthy nations, novel sources of data from the internet and social media are enabling new approaches to social science research and creating new opportunities for public policy. In developing countries, by contrast, fewer sources of such data exist, and researchers and policymakers often rely on data that are unreliable or out of date. Here, we develop a new approach for measuring the dynamic welfare of individuals remotely, and in near real-time, through analyzing their patterns of mobile phone use. To benchmark these methods, we conducted high-frequency panel surveys with 1,200 Afghan citizens, and with the respondent’s consent, matched each individual’s responses to his or her entire history of mobile phone-based communication, which we obtained from Afghanistan’s largest mobile operator. We show that mobile phone data can be used to accurately estimate the social and economic welfare of respondents, and that machine learning models can be used to infer the onset and magnitude of positive and negative shocks. These results have the potential to transform current practices of policy monitoring and impact evaluation.

*University of California at Berkeley, School of Information, jblumenstock@berkeley.edu.
†University of California at San Diego, Rady School of Management, mjcallen@ucsd.edu.
‡Washington University in St Louis, Olin Business School, tghani@wustl.edu.
§University of California at Berkeley, School of Information, nkeleher@berkeley.edu.
¶Princeton University, Woodrow Wilson School, jns@princeton.edu
1 Introduction

Accurate and timely estimates of population characteristics are a critical input to social and economic research and policy. The geographic distribution of poverty and wealth is used to make decisions about resource allocation, and provides a foundation for the study of inequality and the determinants of economic growth (Kuznets, 1955; Fields, 1989). In developing countries, however, the scarcity of reliable quantitative data represents a major challenge to policymakers and researchers. In much of Africa, for instance, national statistics on economic production may be off by as much as 50 percent (Jerven, 2013). Spatially disaggregated data, necessary for small area statistics and used in both private and public sectors, often do not exist (Ghosh and Rao, 1994; Elbers et al., 2003).

In wealthy nations, novel sources of passively-collected data are enabling new approaches to demographic modeling and measurement (Lazer et al., 2009; King, 2011). Data from social media and the “internet of things,” for instance, have been used to measure unemployment (Choi and Varian, 2012; Antenucci et al., 2014), electoral outcomes (Wang et al., 2015), and economic development (Eagle et al., 2010). While most such sources of “big data” are scarce in the world’s poorest nations, the notable exception is the mobile phone, which is used by 3.4 billion individuals worldwide, and which is becoming increasingly ubiquitous in developing regions (GSMA, 2014).

Over the past several years, there have been significant advances in the literature that demonstrate the potential for using large-scale digital data to construct fine-grained estimates of wealth and welfare in developing countries. Early work by Eagle et al. (2010) linked regional patterns of mobile phone use to wealth in developed nations, while more recent research shows how mobile phone data can also construct accurate measures of poverty in developing countries (Blumenstock et al., 2015; Blumenstock, 2014). In parallel, researchers have shown how satellite night-light data (Henderson et al., 2012; Chen and Nordhaus, 2011) and daytime imagery (Jean et al., 2016; Blumenstock, 2016) can be used for small-area wealth estimation.
In this paper, we show that the large volumes of transactional data collected by mobile 
phone companies can not only be used to reconstruct a cross-sectional estimate of the distri-
bution of wealth, but that these high-frequency data can also be used to construct dynamic 
measures of the changing socioeconomic state of the underlying population. In particular, 
we show that the negative and positive economics shocks experienced by individuals – such 
as an unexpected illness in the household, or the receipt of an unexpected gift – are reflected 
in the mobile phone transaction logs of those individuals, and that a machine learning al-
gorithm trained on a small labeled set of respondents can be used to predict the shocks 
experienced by an entire country’s population.

Our analysis is facilitated by the construction of a novel dataset that links high-frequency 
panel survey data to a large administrative database of mobile phone transaction logs. Specif-
ically, we recruited 1,234 Afghan mobile phone subscribers from two provinces in Afghanistan 
to participate in a 7-month panel survey, which consisted of two face-to-face surveys at base-
line and endline, and phone-based surveys every two weeks in the period between face-to-face 
surveys. In each survey round, we collected detailed information about household income, 
expenditures, asset wealth, exposure to shocks, and a host of other socioeconomic variables. 
Each individual’s survey data was then linked to an administrative dataset capturing the 
universe of mobile phone communications on Afghanistan’s primary mobile phone network. 
Informed consent was received from each respondent before these two datasets were merged.

After merging these two datasets, we first replicate results in Blumenstock et al. (2015), 
and show that an individual’s historical logs of mobile phone use can be used to infer his 
wealth. This approach involves first converting each individual’s mobile phone logs - which 
contain tens of thousands of phone calls and text messages - into a set of aggregate metrics 
that quantify patterns of phone use. In total, our algorithm generates thousands of such 
metrics, which capture intuitive statistics such as “total number of calls made per day” as 
well as more obscure properties such as “average entropy of communication of the individual’s 
closest friends.” We then apply several common supervised learning algorithms to determine
which metrics are the best joint predictors of wealth, using cross-validation to help ensure that the model will generalize.

We next show that the social and economic shocks experienced by respondents can be observed in their patterns of mobile phone communications. Over the course of our 7-month study, subjects experienced a large number of negative shocks (including job loss, illness, and death in the household) and positive shocks (such as new jobs, gifts received, and births). These events are directly reported in the survey data, and the impact of these events can also be seen in self-reported data on consumption and expenditures, as well as in more subjective indices of mental health and emotional well-being. Most importantly, we show that such events induce statistically significant changes in the individual’s patterns of mobile phone use.

To more precisely estimate the impact of known income shocks, we conducted a field experiment that experimentally delivered a positive income shock to each respondent, where the timing and magnitude of the shock was experimentally determined. Specifically, each respondent was into one of three treatment conditions: (i) 400 respondents received a small in-kind gift plus a cash gift of 3500 Afs. (roughly 2 weeks salary); (ii) 400 received a small in-kind gift plus a cash gift of 350 Afs. (roughly 1 day’s pay); and (iii) 400 received only the in-kind gift. As with the self-reported shocks captured in the survey data, we see that patterns of phone use change in response to this exogenous income shock, and that the response is different for each treatment group.

Our final set of results indicate that patterns of mobile phone use are sufficient to predict the onset and magnitude of real-life changes in welfare. Specifically, we develop a supervised learning algorithm that can infer the existence of a positive or negative shock from the continuous streams of phone data.

Taken together, these results also suggest a powerful new paradigm for development research and policy in contexts where reliable sources of quanitative data are scarce. We highlight two immediate applications. First, this approach could be used to detect the
onset of negative economic shocks and rapidly target resources to vulnerable populations. Such targeted intervention might offset the far-reaching and long-lasting effects of short-term shocks (Maccini and Yang, 2009; Pissarides, 1992; Bowles et al., 2006; Carter and Barrett, 2006). Second, these methods can engender extremely cost effective methods for policy monitoring and impact evaluation. Indeed, if changes in welfare can be inferred from data already being captured by mobile phone operators, it suggests that program impact could be quickly estimated at a fraction of the cost of traditional methods.

2 Data

To enable this research, we developed a unique dataset that matches panel data from surveys of 1,200 households to administrative data on mobile phone use. Thus, for each of surveyed individuals, we observe detailed information about every phone call and text message in which the individual was involved, continuously over time, for a period of several years. Matching those data to the self-reported survey data, we can develop an understanding of how patterns of communication change in response to real-world events, and build models that use the phone data to infer changes in the welfare of the individual. Before discussing these models, we describe the two separate sources of data.

2.1 High-frequency panel survey

From October 2015 to May 2016, we conducted interviews of 1,200 individuals in two provinces, Kabul and Parwan, of Afghanistan. Individuals were sampled for inclusion in the research using two-staged clustered random sampling. First, 750 enumeration areas were stratified by province and estimated socio-economic status. Within each stratum, a enumeration areas were randomly sampled with replacement using sampling weight proportional to enumeration area population. Within each sampled enumeration area, we endeavored

\footnote{We excluded enumeration areas in 5 districts of Parwan Province due to security concerns.}
to survey one individual from 10 households.\textsuperscript{2} Eligible households were those within the enumeration area that reported ownership of a mobile phone account with a major telecommunications provider in Afghanistan. Eligible households make up approximately two-fifths of all households in the enumeration areas (36.5\% in Kabul Province and 45.5\% in Parwan).

In total, we surveyed 1,200 households through in-person interviews in October 2015. In nearly all cases (99\%), a male member of the household was interviewed. Prior to the interview, surveyors explained that the purpose of the research was to use cellphone records to measure economic activity in Afghanistan. Subjects were asked to provide voluntary consent to participate in the panel survey. Consenting individuals agreed to permit us to combine survey data with cellphone records (date, time, duration, recipient id, and cell tower of calls and texts). Respondents were also informed that they may be selected to participate in bi-weekly phone surveys following the in-person visit. Respondents were provided with a small in-kind gift for participating in the baseline survey.

Appendix Table A1 shows baseline demographic information for households surveyed. Respondents are, on average, 40 years old. Mean household size is 8.18 individuals (median of 7), roughly 5 adults and 3 children, of whom 1.29 are economically active outside of the home. Social contacts are skewed towards local, e.g. within neighborhood, connections. Median per-capita weekly income among sampled households were 444 Afghani (USD$6.88 at the time of the baseline survey wave) and median per-capita expenditures were 400 Afghani (USD$6.20). 12\% of households had cash savings at home and 8\% have savings in a bank account. The mean number of phones per household is 3.4 with slightly more SIM cards (each SIM card corresponds to a unique MSISDN, or phone, number).\textsuperscript{3}

Table A1 also shows that Kabul and Parwan samples are statistically different along many dimensions. Kabul respondents are older, live in larger households, have lived is their neighborhood longer, and report per capita incomes close to three times higher than the

\textsuperscript{2}Because enumeration areas were sampled with replacement, number of households surveyed per enumeration area ranged from 10 to 40.

\textsuperscript{3}The mean number of SIM cards is greater that the number of phones because some phones hold multiple SIM cards and some SIM cards are un-used.
average Parwan respondent.

Following the baseline survey round, we conducted follow-up phone interviews with 915 households over the course of a six-month period. Phone interviews were conducted on a fixed schedule, with enumerators calling sampled respondents every two weeks in order to conduct a short (25 minute) interview. As summarized in Figure 1, 79% of respondents completed at least 11 phone interviews between November 2015 and May 2016. Response rates were similar in the two provinces. At three different times, in-person visits were made to each of household in the phone sample in order to deliver a small, in-kind gift.

Figure 1: Number of phone surveys completed

Notes: A total of 915 respondents were sampled and called on a bi-weekly basis to participate in the high-frequency phone surveys. 50% of respondents completed at least 8 phone interviews and 79% completed at least 11 phone interviews. 118 respondents were never reached during the phone survey rounds.

---

4 An original sample of 800 households, 400 in each province, was randomly sampled from the baseline respondents. After the second month, we increased the sample to 915 by drawing an additional 115 respondents from the remaining baseline respondents to be included in the following phone survey waves.

5 The value of the in-kind gift was approximately USD$2.
2.2 Mobile Phone Metadata

Building on a growing literature that uses mobile phone metadata to analyze social and economic indicators, we partnered with one of Afghanistan’s largest telecommunications providers to combine the panel survey data with high-frequency mobile phone transaction data.\(^6\) Mobile phone Call Detail records (CDR) contain basic metadata on all transactions mediated by the mobile phone network, including phone calls, text messages, and mobile money use. In total, we observe tens of thousands of transactions, each of which contains several fields including: the identity of the calling (or sending, in the case of text messages) and receiving parties, the date and time of the event, the duration and cost of the cost, as well as the location of the cell phone tower nearest to the two parties at the time of the event. Transactions may include calls, texts, data usage, or mobile money transactions.

In much of the analysis that follows, we collapse these logs of continuous transactions at the level of the household. For instance, our canonical dataset is a biweekly panel of aggregate call activity, so that if a phone survey was conducted on March 15, we use all phone transactions from March 1 to March 14 to compute a respondent’s call activity for the survey wave corresponding to the given date. For households with more than one phone number (MSISDN), we compute indicators at the household level. In total, our database of mobile phone transactions includes 2,983 unique MSISDNs for our 1,200 household sample. Appendix Table A2 provides a description of a representative set of the metrics that are computed from the call data.\(^7\) The latitude and longitude of the cellphone tower used by the calling party to initiate a call is used to approximate the location of the caller.

Table 1 presents summary statistics for our phone indicators in the two weeks prior to the baseline face-to-face survey wave. Among our sample, households are active 10 of the preceeding 14 days. Households interact with, on average, 24 unique contacts with 3.25 transactions per contact during the two-week period. On average, households engage in 94

\(^6\)See Onnela et al. (2007); Eagle et al. (2009); Blondel et al. (2015); Blumenstock (2014); Blumenstock et al. (2015) for discussion of common analytical approaches in using mobile phone metadata.

\(^7\)We use a publicly available library to construct these aggregates, see de Montjoye et al. (2016).
calls for a total of 60 minutes during the pre-baseline period. The mean household initiated 51% of their calls, 26% of calls are close to their home, and 19% are during the night hours.

Table 1: Baseline Mobile Phone Indicators

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th>Kabul</th>
<th>Parwan</th>
<th>T-Test (P-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Days</td>
<td>10.08</td>
<td>10.56</td>
<td>9.56</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(4.26)</td>
<td>(4.07)</td>
<td>(4.41)</td>
<td></td>
</tr>
<tr>
<td>Number of Call Contacts</td>
<td>24.31</td>
<td>26.62</td>
<td>21.82</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(26.98)</td>
<td>(28.39)</td>
<td>(25.17)</td>
<td></td>
</tr>
<tr>
<td>Calls Per Contact</td>
<td>3.25</td>
<td>3.29</td>
<td>3.21</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>(2.30)</td>
<td>(2.25)</td>
<td>(2.35)</td>
<td></td>
</tr>
<tr>
<td>Number of Calls</td>
<td>93.92</td>
<td>101.37</td>
<td>85.87</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(131.21)</td>
<td>(129.07)</td>
<td>(133.13)</td>
<td></td>
</tr>
<tr>
<td>Number of Outgoing Calls</td>
<td>74.32</td>
<td>76.05</td>
<td>72.45</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>(69.07)</td>
<td>(71.16)</td>
<td>(66.76)</td>
<td></td>
</tr>
<tr>
<td>Total Call Duration</td>
<td>60.15</td>
<td>65.89</td>
<td>53.94</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(91.47)</td>
<td>(92.99)</td>
<td>(89.46)</td>
<td></td>
</tr>
<tr>
<td>Proportion of Calls Initiated</td>
<td>0.51</td>
<td>0.53</td>
<td>0.50</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.30)</td>
<td>(0.32)</td>
<td></td>
</tr>
<tr>
<td>Proportion of Calls from Home</td>
<td>0.26</td>
<td>0.24</td>
<td>0.28</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.21)</td>
<td>(0.21)</td>
<td></td>
</tr>
<tr>
<td>Proportion Night Time Calls</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.16)</td>
<td>(0.16)</td>
<td></td>
</tr>
<tr>
<td>Number of Cell Towers</td>
<td>20.33</td>
<td>23.21</td>
<td>17.20</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(19.04)</td>
<td>(19.87)</td>
<td>(17.59)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Table reports average aggregate statistics of mobile phone use for all respondents (column 1), and then separately for respondents in Kabul (column 2) and Parwan (column 3). Column 4 indicates the p-value from a two-sided t-test of the difference between Kabul and Parwan. Statistics are reported for the 2-week period prior to baseline. Standard deviations in parentheses.

3  Results

3.1 Measuring welfare from mobile phone metadata

In recent work, we have shown that anonymized data from mobile phone networks can be used to predict the poverty and wealth of individual subscribers, and that these predictions
can be used to create high-resolution maps of the geographic distribution of wealth (Blumenstock et al., 2015). Our modeling approach combines feature engineering with feature selection, first transforming each persons mobile phone transaction logs into a large set of quantitative metrics, then winnowing out metrics that are not predictive of wealth. The first step employs a structured, combinatorial method to automatically generate several thousand metrics from the phone logs that quantify factors such as the total volume, intensity, timing, and directionality of communication; the structure of the individuals contact network; patterns of mobility and migration based on geospatial markers in the data; and so forth. A representative list of features generated through this combinatorial process are shown in Table A2.

The second step uses regularization methods to filter out uninformative phone metrics and select a model that is more likely to generalize. We test several common supervised methods, including a lasso, ridge, and “elastic net” regularization (Zou and Hastie, 2005), as well as common tree-based methods (Breiman et al., 1984), but find little impact of the choice of learner on the overall model performance. In the results below, we show performance from a gradient boosting algorithm (see Friedman, 2001), which compares favorably with other models and can be quickly optimized via gradient descent.

Figure 2 presents results most directly comparable to Blumenstock et al. (2015), showing the model’s ability to predict a composite wealth index (defined as the first principal component of several assets and fixed housing characteristics), measured in the baseline survey, using 6 months of mobile phone metadata prior to baseline. The cross-validated test $R^2$ of 0.375 compares favorably to results in Blumenstock et al. (2015), which achieved a cross-validated test $R^2$ of 0.47.

Superficial attempts to generalize this approach from the cross-section to the panel produce mixed results. For instance, if we use 2 weeks of historical phone use instead of 6 months of data, the accuracy of the predictions is substantially lower. In Figure 3, we show the results of training separate models to predict the total expenditures of respondents in
Figure 2: Performance in predicting total expenditures, pooling data from multiple waves.

In this figure, mobile phone metrics are computed from the two weeks of call data prior to the data of the survey. As can be seen in the figure, the cross-validated $R^2$ is significantly lower for each survey wave than it is for the model estimated in Figure 2. In Section 3.3, we return to this issue and develop a more robust model designed to estimate short-term changes in welfare from limited data on mobile phone use.

### 3.2 The impact of negative and positive shocks

The participants in our study regularly experienced negative and positive shocks. For instance, over the course of the 7-month survey period, 43% of individuals reported that someone in their household was too ill to work, and 38% reported a large rise in food prices. The full extent of experienced shocks, as reported in each wave of the survey data, is shown in Figure 4. These idiosyncratic shocks, as well as the positive income shock delivered ex-
impact experimentally, are the primary source of variation that we seek to detect in the mobile phone data.

**Impact of shocks on self-reported outcomes**

We first provide evidence that the shocks reported by respondents have meaningful consequences on their economic livelihoods. Our starting point is the analysis of the randomized cash transfer, as this shock was unequivocally exogenous. To measure the impact of this
shock, we run regressions of the form:

\[ Y_{it} = \alpha + \beta_1 \text{Large}_{it} + \beta_2 \text{Small}_{it} + \mu_i + \pi_t + \epsilon_{it} \]  

where \( Y_{it} \) indicates the welfare outcome of individual \( i \) in survey wave \( t \), and \( \mu_i \) and \( \pi_t \) are individual and survey-wave fixed effect. \( \text{Large}_{it} \) and \( \text{Small}_{it} \) are dummy variables indicating whether the individual received the large (3500 Afs.) or small (350 Afs.) cash transfer in wave \( t \), with the omitted category being individuals who received an in-kind gift. The primary coefficients of interest are \( \beta_1 \) and \( \beta_2 \), which indicate the causal effect of the income shock on the welfare outcome of interest.

We separately estimate the impact of the income shock on four key welfare outcomes: (i) Food expenditures; (ii) total expenditures; (iii) total savings; and (iv) the wealth index – see Section 2.1 for more details on how these variables are constructed. Results are shown in Figure 5. While the small cash transfer had only a modest impact on expenditures, the large cash transfer increased food and total expenditures by roughly 0.3 standard deviations,
Figure 5: The impact of the randomized cash transfer on self-reported welfare outcomes and savings by 0.18 standard deviations. Neither cash transfer had an immediate effect on the composite wealth index, which is not surprising since the index is based on assets, which likely take longer to accrue.

Other self-reported shocks have similarly large effects on consumption and expenditures. Using a specification similar to (1), we separately estimate the impact of several of the most commonly experienced shocks on total food expenditures, and plot the coefficients in Figure 6. The first two rows replicate earlier results, while the latter five rows show how a variety of experienced shocks impact expenditures. For this figure “electricity outage” indicates a major and unpredictable failure of the power grid in Kabul that occurred during our study.\(^8\)

\(^8\)See http://www.nytimes.com/2016/02/18/world/asia/afghanistan-hardship-taliban-bombings.html
Figure 6: The impact of self-reported shocks on expenditures
Impact of shocks on mobile phone metadata

Our next set of results test the extent to which economic shocks are manifest in the way that individuals use their phones. Prior descriptive work on this topic shows that major real-world events have a striking impact on call activity (Bagrow et al., 2011; Altshuler et al., 2013; Gao et al., 2014; Dobra et al., 2015). For example, Figure 7 shows the immediate impact of an earthquake in Rwanda on the calling behaviors of those affected. Immediately after the earthquake there is a sharp spike in calls between people affected by the quake, followed by a spike in calls from people inside the affected region to outside, followed by a more gradual increase in calls from outside the epicenter to inside. Our empirical question is whether smaller, idiosyncratic shocks exhibit similar signatures.

Figure 7: Communication response following an earthquake

Notes: Immediately following an earthquake in Rwanda, there is a spike in calls between people who live near the epicenter, followed shortly thereafter by an increase in calls by those affected to people who were not affected. Over the next hour, there is a steady increase in calls from people living outside to people living near the epicenter. Communication activity between people living far from the epicenter is not affected by the earthquake.

We begin by using model (1) to estimate the effect of positive and negative shocks on a handful of canonical metrics of mobile phone use. Figure 8 shows the response of mobile phone use to the experimental cash transfer, and Figure 9 shows how one dimension of mobile phone use – the number of unique contacts with whom the subject communicates – changes in response to a large number of self-reported shocks.

These figures provide suggestive evidence that following positive and negative shocks, people change their pattern of mobile phone communication. For instance, people increase the duration (in seconds) spent on the phone following both cash transfers (Figure 8). Why
the increase is larger for the small cash transfer is not immediately clear. One hypothesis supported by our survey data is that the small gift of cash increases expenditures on consumption goods, while recipients of the large cash transfer keep the money for larger purchases and savings. It is similarly not obvious why people would communicate with a smaller number of unique contacts after the cash transfers (Figure 9), although such a result would be consistent with a desire to shirk from social obligations (cf. Jakiela and Ozier, 2015). The impact of celebrations is more clear-cut: when families celebrate, the number of unique people with whom they communicate increases by roughly 0.08 standard deviations.

### 3.3 Estimating shocks and changes in welfare from mobile phone metadata

Thus far, we have shown that positive and negative shocks are common in the Afghan population we study; that these shocks have real economic consequences; and that people alter
Figure 9: Impact of shocks on the number of people with whom the respondent communicates their mobile phone use in response to shocks. Next, we develop an approach to modeling changes in welfare from dynamic patterns of mobile phone use. We start from the observation that a model which is able to predict asset-based wealth in the cross section does not trivially generalize to predicting changes in wealth in the panel, or to predicting a more comprehensive set of welfare outcomes. The first point was previewed in Section 3.1, where we showed that the accuracy of a model trained on 6 months of data degraded considerably when trained on only 2 weeks of data. The problem, however, is more fundamental: we have no reason to assume that the relationship between mobile phone use and wealth will be stable over time.⁹

Indeed, we find that a model trained to predict wealth from phone use at time $t$ performs quite poorly in predicting wealth at $t \pm k$. Figure 10 shows how a model trained in wave 2 performs in adjacent waves. To construct this figure, we used the same cross-validated

---

⁹In other contexts, this erroneous assumption has had severe consequences (Lazer et al., 2014)
Figure 10: Predictive performance of model over time

Notes: We train a model to predict wealth using survey data and mobile phone data from wave 2. We then test the ability of this model to accurately predict wealth from mobile phone data in other waves of the survey.
gradient boosting algorithm used to construct Figure 2, but trained it to predict total expenditures in wave 2 using the two weeks of mobile phone metadata prior to baseline. We then hold the model fixed and iteratively update the input data from $t = 0$ until $t = 13$, and show the $R^2$ of the predicted $Y_{i,t+k}$ relative to the true $Y_{i,t+k}$ reported in wave $t+k$. Benchmark performance in wave 2 is quite high, with $R^2 = 0.52$. In waves 1 and 3 – two weeks before and after wave 2, performance degrades to approximately $R^2 = 0.24$, and any farther from wave 2 the performance is closer to $R^2 = 0.13$.

The inability of the baseline model to trivially generalize across periods is due, at least in part, to the changing nature of the relationship between phone use and wealth. Figure 11 highlights the differences in the models trained on successive waves of data by plotting the variable importances of features selected by both models. The variable importance provides an indication of the relative extent to which each feature influences the predictions of the model (see Friedman, 2001). While both models share certain commonalities (e.g., the most frequent weekday antenna is quite important to both models), there are significant differences, even though the models were trained just two weeks apart.

4 Discussion and Conclusions

Our results to date validate recent work indicating the mobile phone data can be used to infer socioeconomic characteristics of populations in developing countries. While the Afghan context of this study is quite different from the Rwandan context of prior work (Blumenstock et al., 2015), we find that with only a few weeks’ of mobile phone data, it is possible to accurately infer both the wealth and the expenditures of an individual.

We also describe initial attempts to generalize this approach from cross-sectional wealth estimation to a non-stationary regime. Initial results here indicate that positive and negative shocks are common in the Afghan population we study; that these shocks have real economic consequences which are highlighted in our survey data; and that people alter their mobile
Figure 11: Comparison of feature importances for model trained on two successive waves of surveys

Notes: The scaled variable importance indicates the relative extent to which each feature influences the predictions of the model.
phone use in response to shocks. Taken together, these results suggest that mobile phone
data can likely be used to detect economic shocks, and to estimate changes in welfare over
time.

In ongoing work, we are developing more sophisticated econometric and machine learning
models to predict changes in well-being from changes in the continuous, real-time stream of
mobile phone transactions data. As we have shown in this paper, this is not a trivial modeling
exercise, since naive “cross-sectionalized” models do not generalize well over time. This
appears to be due to the fact that the underlying relationship between phone use and wealth
is changing over time. Nonetheless, we remain optimistic that there exists an appropriate and
more nuanced approach to modeling the dynamic relationship between welfare and phone
use.

If such methods prove successful, the implications for development research and policy-
making are profound. We highlight two potential applications. First, this approach could be
used to detect the onset of negative economic shocks and rapidly target resources to vulner-
able populations. Such targeted intervention might offset the far-reaching and long-lasting
effects of short-term shocks (Maccini and Yang, 2009; Pissarides, 1992; Bowles et al., 2006;
Carter and Barrett, 2006). Second, these methods can engender extremely cost effective
methods for policy monitoring and impact evaluation. Indeed, if changes in welfare can
be inferred from data already being captured by mobile phone operators, it suggests that
program impact could be quickly estimated at a fraction of the cost of traditional methods.
Appendix Tables
### Table A1: Baseline Demographics and Income

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th>Kabul</th>
<th>Parwan</th>
<th>T-Test (P-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kabul</td>
<td>0.51</td>
<td>1.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Male Respondent</td>
<td>0.99</td>
<td>0.98</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Age of Respondent</td>
<td>39.67</td>
<td>41.46</td>
<td>37.78</td>
<td>0.00</td>
</tr>
<tr>
<td>Adult Males</td>
<td>2.68</td>
<td>2.85</td>
<td>2.51</td>
<td>0.00</td>
</tr>
<tr>
<td>Adult Females</td>
<td>2.48</td>
<td>2.63</td>
<td>2.33</td>
<td>0.00</td>
</tr>
<tr>
<td>Male Children</td>
<td>1.54</td>
<td>1.64</td>
<td>1.43</td>
<td>0.02</td>
</tr>
<tr>
<td>Female Children</td>
<td>1.51</td>
<td>1.59</td>
<td>1.42</td>
<td>0.06</td>
</tr>
<tr>
<td>Female/Male Ratio</td>
<td>1.04</td>
<td>0.97</td>
<td>1.11</td>
<td>0.00</td>
</tr>
<tr>
<td>Years in current location</td>
<td>23.71</td>
<td>17.92</td>
<td>29.90</td>
<td>0.00</td>
</tr>
<tr>
<td>Total Contacts - Inside</td>
<td>88.01</td>
<td>96.45</td>
<td>78.97</td>
<td>0.01</td>
</tr>
<tr>
<td>Total Contacts - Outside</td>
<td>67.88</td>
<td>68.12</td>
<td>67.61</td>
<td>0.93</td>
</tr>
<tr>
<td>Number of Phones</td>
<td>3.44</td>
<td>3.94</td>
<td>2.90</td>
<td>0.00</td>
</tr>
<tr>
<td>Number of SIM Cards</td>
<td>3.97</td>
<td>4.58</td>
<td>3.32</td>
<td>0.00</td>
</tr>
<tr>
<td>Working Household Members</td>
<td>1.29</td>
<td>1.41</td>
<td>1.16</td>
<td>0.00</td>
</tr>
<tr>
<td>Per Capita Income (Afs)</td>
<td>1,053.55</td>
<td>1,550.15</td>
<td>527.78</td>
<td>0.00</td>
</tr>
<tr>
<td>Per Capita Expenditures (Afs)</td>
<td>619.73</td>
<td>806.02</td>
<td>422.49</td>
<td>0.00</td>
</tr>
<tr>
<td>Has Savings at Home</td>
<td>0.12</td>
<td>0.06</td>
<td>0.19</td>
<td>0.00</td>
</tr>
<tr>
<td>Has Savings in Bank</td>
<td>0.08</td>
<td>0.10</td>
<td>0.06</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Standard deviations in parentheses.
Table A2: Representative features of mobile phone use

<table>
<thead>
<tr>
<th>CDR variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active days</td>
<td>The number of days during which the user was active</td>
</tr>
<tr>
<td>Number of contacts</td>
<td>The number of contacts the user interacted with</td>
</tr>
<tr>
<td>Calls per contact</td>
<td>The number of interactions a user had with each of its contacts</td>
</tr>
<tr>
<td>Number of calls</td>
<td>The number of interactions, incoming and outgoing</td>
</tr>
<tr>
<td>Call duration</td>
<td>The duration of the user’s calls</td>
</tr>
<tr>
<td>Proportion of Calls Initiated</td>
<td>The percentage of calls initiated by the user</td>
</tr>
<tr>
<td>Proportion of Calls from Home</td>
<td>The percentage of interactions the user had while at home</td>
</tr>
<tr>
<td>Proportion Night Time Calls</td>
<td>The percentage of interactions between 7pm and 7am</td>
</tr>
<tr>
<td>Number of Cell Towers</td>
<td>The number of unique cell towers used</td>
</tr>
</tbody>
</table>
References


Dobra, Adrian, Nathalie E. Williams, and Nathan Eagle, “Spatiotemporal detection of unusual human population behavior using mobile phone data,” *PloS one*, 2015, 10 (3), e0120449.


Lazer, David, Alex (Sandy) Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James


