How Do the Poverty Scorecard and the PAT Differ?

Mark Schreiner

1 March 2014

This document is available at `microfinance.com/#Poverty_Scoring`

Abstract

The poverty scorecard and the Poverty Assessment Tool (PAT) are simple, low-cost ways to measure households' poverty. How do the two differ? For estimating a group's poverty rate, both are unbiased, and the scorecard has smaller standard errors. For targeting individual households, the PAT correctly classifies about one more household per 100. The scorecard has an edge in availability, recentness, and transparency.

Acknowledgements

The Ford Foundation funded this work but did not mandate the topic and is not responsible for the content. Thanks go to Frank Ballard, Dean Caire, Frank DeGiovanni, Sean Kline, Mary Jo Kochendorfer, Anthony Leegwater, Margaret Richards, Jeff Toohig, and Matt Walsh. The PAT was developed for USAID at the University of Maryland by the now-defunct IRIS Center. The author developed the poverty scorecard with support from Grameen Foundation (GF) and the CGAP/Ford Social Indicators Project. The poverty scorecard is the same as what GF calls the PPI®, a performance-management tool that GF promotes to help organizations achieve their social objectives more effectively.

Author

Mark Schreiner is the developer of the poverty scorecard. He directs Microfinance Risk Management, L.L.C., and he is a Senior Scholar at the Center for Social Development at Washington University in Saint Louis.

How Do the Poverty Scorecard and the PAT Differ?

1. Introduction

The poverty scorecard and the Poverty Assessment Tool (PAT) are simple, low-cost ways for pro-poor organizations to measure their participants' poverty. The tools can estimate the share of participants with expenditure below a poverty line, and they also can be used to target services based on poverty. They seek to bolster accountability and to make social-performance management more quantitative and intentional.[1]

This paper compares the scorecard and the PAT. The scorecard is developed by the author of this paper and sponsored by the Ford Foundation/CGAP Social Indicators Project and by the Grameen Foundation, while the PAT was developed by IRIS Center for USAID. Both the scorecard and the PAT are simple and low-cost. While the two tools take different statistical approaches, they have similar accuracy. In particular, both are unbiased when estimating poverty rates (the scorecard has smaller standard errors), and they have similar targeting accuracy (the PAT correctly classifies about one more household per 100). The scorecard is available in more countries than

[1] Poverty scorecards ("scorecards" for short) are also called "simple poverty scorecards", "Progress out of Poverty Indexes®", or "PPIs®" (trademarks registered by Grameen Foundation). All the names refer to the same approach, and most scorecards are also branded as PPIs®. Like PATs (`povertytools.org`), scorecards are available at no cost (`progressoutofpoverty.org` or `microfinance.com`). Copyright in a given scorecard is held by its sponsor and by Microfinance Risk Management, L.L.C.

the PAT (53 versus 36). In the 14 countries that have a scorecard and a PAT derived from data of different vintages, the scorecard uses more recent data in all but one.

In countries that have a scorecard and a PAT derived from the same data, the comparison must hinge on something other than accuracy, availability, and recency. In particular, the scorecard's edge in transparency helps non-specialists to understand its potential usefulness and increases its up-take.

For example, the scorecard's simple formula for linking indicators to estimates of poverty rates is open and shown to users, but the PAT's less-simple formula is not revealed to users. The scorecard shows how a household's indicators correspond with points, how points are added up to get a score, how scores are converted with a look-up table to probabilities of being below a poverty line, and how probabilities are averaged across households to estimate a poverty rate. This transparency builds trust and boosts up-take.[2]

The scorecard is also better documented. For example, its poverty lines and their sources are reported more completely. For estimating poverty rates, the scorecard's accuracy is described in terms of bias and standard errors. For targeting, the scorecard reveals households' scores, tells how to use scores for targeting, and reports targeting

[2] The world's first scoring tools were likewise simple and transparent. "Because the first Fair, Isaac credit-scoring systems were to be deployed in small towns in rural America at the point of sale, they had to be simple enough to be understood by people with no knowledge of statistics and no access to calculators. [This drove] the choice of statistical method as well as the card format. . . . Scoring was to be done manually by retail clerks, [so] addition . . . was possible but multiplication [was not]" (Poon, 2007, p. 289).

accuracy for a range of cut-offs. In contrast, the PAT hides households' scores, reports targeting accuracy for only one cut-off, and advises against targeting individual households.[3] While neither tool is appropriate in all contexts and for all purposes, the scorecard helps potential users to make an informed decision.

Transparency differs in part because the two tools were born with different goals. USAID created the PAT to monitor compliance with a law that requires that "50 percent of all microenterprise resources be targeted to clients who are 'very poor'" and that directs USAID to report "the percentage of assistance . . . allocated to the 'very poor'" (U.S. Congress, 2004). The PAT inherited a "proving" purpose and a guaranteed user base.

The scorecard approach grew out of a home-grown tool made by a pro-poor microlender in Bosnia-Herzegovina that wanted to understand its clients better (Matul and Kline, 2003; Schreiner *et al.*, 2004).[4] From the school of hard knocks, the scorecard's developer—that is, the author of this paper—learned that, regardless of accuracy, complex tools do not get used, and the scoring literature teaches that simple tools can be about as accurate as complex ones. Armed with these lessons and lacking a guaranteed user base, the scorecard was designed to attract voluntary adoption via low cost, ease-of-use, and transparency. Rather than help users *prove* their poverty

[3] FHI360 (2013, p. 9); `povertytools.org/povertypres/USAID_PATs/player.html` (slide 24), retrieved 26 February 2014.
[4] Nothing about the scorecard or PAT is specific to microfinance/microenterprise, and both tools are used in other sectors such as health, education, and agriculture.

outreach, the scorecard aims to help them *improve* their social-performance management.

Of course, both the scorecard and the PAT can serve for internal "improving" purposes or external "proving" mandates. Indeed, two scorecards are branded as PATs,[5] and some organizations use the PAT voluntarily. But the scorecard's focus on management as opposed to monitoring—and the imperative to earn voluntary adoption—led to greater incentives for transparency.

The next section compares the scorecard and the PAT in terms of availability, recency, transparency, and construction. The section after that compares accuracy. The last section sums up, looks at the take-up of the two tools, and argues that USAID is not measuring poverty as mandated.

The appendix documents that mistakes in the implementation of poverty lines are disconcertingly common but that the scorecard makes fewer than the PAT and fewer than the World Bank's PovcalNet.

[5] Since 2012, USAID partners have been authorized to use PPI®-branded scorecards.

2. Non-accuracy comparisons

This section compares the scorecard and the PAT in terms of availability, recency, transparency, and construction. The scorecard is more available and more recent, and its construction is more transparent.

2.1 Why use a scorecard or a PAT?

Indirect estimates of poverty via the scorecard or PAT can be useful when direct measurement is costly because many households work in the informal sector or are subsistence farmers. In this case, much cash income goes unrecorded, and much income is not in cash at all but rather in the value of food grown and eaten at home.

An alternative to measuring income (the flow of resources into a household) is to measure consumption expenditure (the flow of resources used up by a household). This means asking about quantities and prices (actual or hypothetical) for hundreds of consumption items in interviews that can last as long as two days. Surveyed households may also be obliged to track their consumption in a diary for weeks.

Expenditure surveys are the most accurate way to measure poverty,[6] but they are too costly for pro-poor organizations that want to take frequent measurements for

[6] A household is said to be *poor* if its expenditure—considering the number of its members and perhaps their age and sex—is below a poverty line such as a country's official national line or the World Bank's international benchmark of $1.25/day at 2005 purchase-power parity. While this expenditure-based definition is not the only (nor the

many participants. Governments do field expenditure surveys, but these may be infrequent, and sometimes it takes years to crunch the data and to report poverty rates.

To make frequent, widespread poverty-measurement feasible, the scorecard and the PAT strike a compromise between accuracy and cost. Given data from a country's national expenditure survey,[7] a tool's developer uses statistics and judgment to select a few simple, verifiable indicators that are strongly correlated with expenditure-based poverty. Common examples of poverty indicators are the number of household members, the ownership of a refrigerator or a television, the main type of cooking fuel, and the type of roof, walls, or floor.

A household's indicators can be collected in 10 to 20 minutes by third-party enumerators or by low-level employees of pro-poor organizations (such as microloan officers, agricultural-extension agents, or health-outreach workers). Each response is linked with a point value (derived from national-survey data), and the sum of the points is the score. The scorecard and PAT then convert the score to an indirect *estimate*—as opposed to a direct *measure*—of the probability that a household has expenditure below a poverty line. These probabilities are averaged across households to estimate a group's poverty rate. If desired, the score can be used for targeting.

---

best) definition of poverty, it is what people and governments usually use, and it sums up poverty in a single, understandable number.

[7] In this paper, *expenditure* is shorthand for "expenditure or income", as some less-poor countries—especially in Latin America—define and measure poverty in terms of income.

2.2 Availability and recency

Table 1 lists the 63 countries with a scorecard or PAT, as well as the year of the survey data used to construct a tool. The scorecard is available in more countries (53 versus 36). There are 27 countries with only a scorecard (but no PAT), 10 with only a PAT (but no scorecard), and 26 with both.

Among the 26 countries with both a scorecard and a PAT, the most-recent version of the tools use the same data in 12 countries. In 13 of the 14 countries where the tools use data of different vintages, the scorecard's data is more recent.

In sum, the scorecard is available for more countries and—when data vintage differs for a country that has both tools—the scorecard is almost always more recent.[8]

2.3 Structure and presentation

Figures 1 and 2 are the paper instruments for Peru for the PAT and the scorecard.[9] Both tools record responses to poverty indicators as well as administrative identifiers, and both tell the enumerator how to collect items in the same way as in the country's expenditure survey. The PAT's instructions and the scorecard's back-page worksheet are integral parts of the tools, as are the PAT's "User Guide" and the scorecard's "Guide to the Interpretation of Scorecard Indicators".

[8] Scorecards are still being made and updated. No PATs have been made or updated since 2011.

[9] Figure 1 is copied from `povertytools.org/countries/Peru/USAID_PAT_PERU_05-2013.xls`, retrieved 3 January 2014. Figure 2 is from Schreiner (2012).

How do the two Peru tools compare in terms of structure and presentation? Length is similar; the PAT (Figure 1) records seven administrative fields (versus nine for the scorecard), asks six questions for each household member (versus three), and has 12 household-level questions (versus 10).

The PAT has two items (urban/rural status and region) that scorecards generally omit. In national surveys, urban/rural status is usually recorded based on census maps before going to the field. Users of a scorecard or PAT, however, are unlikely to consult census maps, even if they are available. Urban/rural status is often obvious, but it may not be, as in peri-urban areas. Scorecards omit urban/rural status because pro-poor organizations are not likely to determine it in the same way as it was determined in the national-survey data used to construct the tool.

Recording a household's region of residence is more straightforward. Still, some users balk at assigning points based on state, department, or province. This may reflect political concerns, it may seem unfair to link location to estimated poverty, or it may be that a region's points—after considering other indicators in the scorecard—do not match intuition.

Regional indicators are feasible for the PAT for two reasons. First, the PAT's paper instrument does not reveal any indicator's points. Second, the PAT does not reveal its indicators, so users may not realize that the region—recorded with administrative identifiers—is also used as a poverty indicator.[10]

The PAT hides its points and formula not to mislead users but rather to discourage cheating. After all, if a tool is used to *prove* poverty outreach—the PAT's purpose for USAID's microenterprise partners—and if the tool is applied by the same entity that is being monitored, then there are incentives to cook data.[11]

The PAT's lower transparency probably does not discourage cheating much. Even if points are hidden, a person with common sense can figure out, for example, that not owning a refrigerator signals more poverty than owning one, or that having finished only grade school signals more poverty than having finished high school.

[10] The PAT's indicators and points are documented, and FHI360 (2013) is a high-level guide to linking responses, points, scores, and poverty estimates. But the process is not made clear by the PAT's paper instrument, and few users—especially front-line workers—will dig up the documentation or understand how the parts fit together.

[11] If the goal is to *improve* internal management, then cheating is self-destructive. Still, perverse incentives crop up with both the scorecard and PAT whenever the entity that measures poverty is also rewarded for finding more poverty. If a tool is used for targeting, then respondents also have incentives to try to look poorer than they are. Neihaus *et al.* (2013) and Drèze and Khera (2010) suggest that one way to deter corruption is to use simple targeting tools.

The PAT's lower transparency also makes it more difficult to see how it works. Asking users for blind trust discourages voluntary adoption and also makes it more likely—if a tool is adopted—that it is used not only for *proving* performance but also for *improving* it.

The scorecard's structure and presentation focus on transparency (Figure 2). The goal is to help non-specialists see how it works so that they can adopt it voluntarily, then apply it properly, and finally act on its results.

To this end, the scorecard has 10 multiple-choice indicators. Each response has a non-negative, whole-number point value.[12] Agents in the field can add up points by hand to get a household's score, which ranges from 0 (highest poverty likelihood) to 100 (lowest poverty likelihood). The paper instrument shows how the score is based straightforwardly on responses to common-sense poverty indicators.

Like the scorecard, the PAT converts responses into points that are added up to get a score. The PAT differs in that its paper instrument does not show clearly what all the indicators are, it does not display the points, and it does not show its math.

Figure 3 presents the Peru PAT (IRIS Center, 2011c) as if it were a scorecard. On the one hand, this reveals the two tools' basic structural similarity. On the other hand, it highlights that the tools differ in presentation/transparency and that the PAT requires dividing, multiplying, adding, subtracting, and squaring point values that have four decimal places and that are often negative. Of course, the math is easily handled

[12] For a given indicator, the point value for the most-likely-poor response is always zero.

with spreadsheets or software, but the point is that voluntary adoption is facilitated if users can see and understand how scores are found.

2.4 From scores, to poverty likelihoods, to poverty estimates

The scorecard converts scores to poverty likelihoods more transparently than the PAT. The scorecard avoids math altogether, using a look-up table to link a score with a likelihood (that is, a probability) that expenditure is below a poverty line. In the case of Peru (Table 2), scores of 25–29 (for example) are linked with poverty likelihoods of 77.0 percent for the national poverty line and of 1.9 percent for $1.25/day 2005 PPP poverty line. A group's estimated poverty rate is then the average of its households' poverty likelihoods.[13]

Where do poverty likelihoods come from? This, too, is simple and transparent. After a scorecard is constructed with data from a country's national expenditure survey, the scorecard is applied to get scores for the same households whose data was used in construction. A given score's poverty likelihood is then derived as the share of surveyed households who have that score and who also have directly measured expenditure below a poverty line. With this process, scores from a single scorecard can be converted to poverty likelihoods for any number of poverty lines.[14]

[13] Scores are converted to poverty likelihoods first, then poverty rates are estimated as the average likelihood. A common mistake is to average scores first and then to estimate the poverty rate as the poverty likelihood associated with the average score.
[14] On average, the scorecard supports eight poverty lines, the PAT three (Table 1).

In the case of Peru, 4,723 (weighted) households in the construction data have scores of 25–29, and 3,635 of these have expenditure below the national poverty line. The poverty likelihood for scores of 25–29 for this line is then 3,635 ÷ 4,723 = 77.0 percent. When a scorecard is applied to a household and the score is 25–29, then that household's poverty likelihood is 77.0 percent because this is the share of households in the national expenditure survey who had this score and who also were poor.

In 33 of the 36 PATs, a score is converted to a poverty likelihood of 100 percent if its inverse logarithm (that is, 2.71828^{score}) is below a poverty line, and to 0 percent otherwise (FHI360, 2013). Like the scorecard, the PAT then estimates poverty rates as the average of the poverty likelihoods of the households in a group.

The PAT's 2.71828^{score} is less transparent than the scorecard's look-up table. Most non-specialist users do not understand this expression (if they can find it), so for them the source of PAT estimates is obscure, damping voluntary adoption.

2.5 Selecting indicators and deriving points

The scorecard and PAT use different statistical approaches to select indicators and derive points. Both tools document the construction process at a high level, as most non-specialists lack the background to understand the technical details. Both tools:

- Apply statistics to data from a country's national expenditure survey in order to:
  - Identify indicators strongly correlated with expenditure-based poverty
  - Derive points so as to give lower scores to households with higher poverty likelihoods (scorecards) or lower expenditure (PATs)
- Focus the statistics on households close to the poverty line used for construction
- Use the judgment of developers and feedback from users to weed out indicators that—regardless of statistical quality—may have unstable relationships with poverty over time or that may be non-intuitive to users, costly to collect, difficult to verify, insensitive to changes in poverty, or non-applicable in some regions of a country

The next two sub-sections discuss the construction of scorecards and PATs. It assumes advanced statistical knowledge, so some readers may want to skip to Section 3.

2.5.1 Scorecard construction

The first step in scorecard construction is to divide households in the national expenditure survey into a *construction sample* (used to select indicators and points and to calibrate scores with poverty likelihoods) and a *validation sample* (used to measure accuracy with data that has not also been used in construction).

Let the n_c households in the construction sample by indexed by i. Given a poverty line, a scorecard's points are based on a vector of coefficients β from a Logit regression that relates a household's directly measured expenditure-based poverty status y_i with the household's vector X_i of responses to survey indicators. Given that y_i is 1 if poor and 0 if non-poor, Logit finds β to maximize:

$$\sum_{i=1}^{n_c} \ln\left(\frac{2.71828^{(2\cdot y_i - 1)\cdot \beta X_i}}{1 - 2.71828^{(2\cdot y_i - 1)\cdot \beta X_i}}\right).$$

The coefficients β have many decimal places and are sometimes negative. The scorecard's simpler point scheme is derived from β as follows (Schreiner, 2010c):

- Find, for each indicator and for the Logit's constant term, the "shifted coefficients" γ by subtracting the minimum β coefficient among the responses for an indicator from each of the β coefficients associated with responses for that indicator
- Find the maximum possible value of γX
- Multiply γ by 100, divide by γX, and round each element to the nearest integer[15]

This gives non-negative, whole-number points in which each indicator's "most-poor" response gets zero points and in which scores range from 0 to 100. The transformation is affine/homothetic, so the score based on the transformed points rank-orders households the same as does the original βX.

Scorecard indicators are selected based on statistics and judgment. First, Logit is used to build a single-indicator scorecard for each candidate indicator. One of these single-indicator scorecards is then selected based on accuracy, acceptability to users

[15] If the highest possible score is not 100 due to rounding, then make it 100 by nudging the points for the rarest "least-poor" response up or down by one or two. This hardly changes ranks and spares users from puzzling about something that matters little.

(based on simplicity, cost of collection, and "face validity" in terms of experience, theory, and common sense), sensitivity to changes in poverty, variety among indicators, applicability across regions, relevance for distinguishing among households just above or below the poverty line, and verifiability (Schreiner *et al.*, 2004; Zeller, 2004).

A series of two-indicator scorecards are then built, each based on the best single-indicator scorecard, with a second candidate indicator added. The best two-indicator scorecard is then selected, again based on statistical and non-statistical criteria. These steps are repeated until the scorecard has ten indicators that work well together.[16]

This selection process is like the common R^2-based stepwise algorithm. It differs from naïve stepwise regression—sometimes called "unwise" regression[17]—in two ways. First, it considers not only statistical accuracy but also user acceptance. Second, statistical power is not taken as the "significance" of the coefficients β nor the "goodness-of-fit" of the regressions. These measures are relevant for the inferential paradigm that is taught in most statistics courses. The goal of a poverty-measurement tool, however, is not to use the estimated β to infer how indicators X affect poverty but rather to estimate the likelihood that a household has expenditure below a poverty line based on the values of its indicators X and the coefficients β.

---

[16] Also, potential scorecard users usually review and field-test 15–20 finalist indicators.

[17] See #12 on STAT-L FAQ, `www-personal.umich.edu/~dronis/statfaq.htm`, retrieved 5 January 2014. According to Ira Bernstein, "Stepwise is no substitute for understanding the statistics, the data, and the domain. In general, because overfitting is a real issue, using theory and diagnostics to choose [indicators] that are somehow 'non-optimal' on the current data can nonetheless produce models that generalize better (and . . . are easier to explain to lay people)."

The prediction paradigm aims to give lower scores to poorer households. A scorecard's power to rank-order is taken as the share of randomly paired households—one poor and one non-poor—in which the poor household has a lower score.[18]

2.5.2 PAT construction

In 17 of 36 PAT countries, construction starts by dividing survey data into construction and validation samples.[19] The PAT then tests one-step and two-step versions of four types of regressions:[20]

- Least-squares to estimate the poverty likelihood
- Probit to estimate the poverty likelihood
- Least-squares to estimate the logarithm of expenditure
- Quantile to estimate the logarithm of expenditure

Of the 36 PATs, 30 end up using one-step quantile regression. The quantile is selected to give unbiased estimates of poverty rates. Also, the targeting accuracy of the quantile regressions is generally a little better than that of alternatives. The two-step quantile is often more accurate than the one-step quantile, but the one-step quantile is preferred because it is simpler.

[18] This is the "concordance index" (SAS Institute, 2004) or the "area under the curve" (AUC) that plots the share of all poor households who have scores below a score percentile (vertical axis) against the score percentile of all households (horizontal axis). This is like a Lorenz curve, with "share of all poor households" replacing "share of total income" and "score percentile" replacing "income percentile". In this sense, AUC is like a Gini coefficient.

[19] The other 19 PATs use all the data for construction.

[20] See IRIS Center (2005) or—for Peru—Zeller, Alcaraz V., and Johannsen (2005).

To derive the PAT's points with quantile regression, let the n_c households in the construction sample be indexed by i, and let X_i be the vector of a household's responses to survey indicators. Given a quantile q,[21] let y_i be the logarithm of a household's expenditure, and let y_q be the q^{th} quantile of the y_i's for all n_c households. Then the PAT's points are the vector of coefficients β that minimizes:

$$\sum_{i=1}^{n_c} \gamma_i \cdot (y_i - \beta X_i),$$

where $\gamma_i = q$ if $y_i \geq y_q$, and
$\gamma_i = q - 1$ if $y_i < y_q$.

Finally, repeat these steps for many quantiles q, selecting the one that makes the estimated poverty rate in the construction sample match the actual poverty rate.

The PAT estimates a household's poverty likelihood as either 100 percent (if $2.71828^{\beta X_i}$ is below a poverty line) or 0 percent. The PAT then averages these poverty likelihoods across households to get an unbiased estimate of the poverty rate.

The PAT's developers select poverty indicators in three stages. First, they identify candidate indicators in the national expenditure survey that have common-sense links with poverty, are inexpensive to collect, and are acceptable to users. Second, they whittle down this pool of practical candidates with an R^2-based stepwise algorithm applied to least-squares regression on the logarithm of expenditure. Finally, they plug 15–18 of the stepwise-selected indicators into quantile regression.

[21] q is between 0 and 1. For example, 0.5 is the median, and 0.25 is the first quartile.

3. Accuracy comparisons

Accuracy is a toss-up. For estimating poverty rates, both the scorecard and the PAT are unbiased, with the scorecard having smaller standard errors. For targeting, the PAT gets about one more household right per 100.

3.1 Comparing apples to apples

The comparisons here are as clean as possible, but they are not perfect. To start with, only 26 of 63 countries have both a scorecard and a PAT. In 11 of these 26, direct comparisons are not possible because the tools are validated against data of different vintages. Five of the remaining 15 countries cannot be compared—even though both tools use data from the same national survey—due to mistakes in the PAT's data handling (Cambodia and Tanzania, see appendix) or due to the PAT's not reporting out-of-sample/in-time measures of accuracy (Ethiopia, Guatemala, and Nepal).

In the remaining 10 directly comparable countries, the comparison is still imperfect because the scorecard and the PAT differ in the size and composition of their construction and validation samples. Also, the two tools are not constructed based on the same poverty line in any of the 10 countries. To be relevant for local users, the scorecard is constructed with the national line in 44 of 53 countries. Following USAID's

definition of "very poor", the PAT is constructed with the highest of the median line or $1.25/day.[22]

Comparisons are still possible because scorecards are calibrated to the median line and to $1.25/day. The poverty rates reported by the scorecard and by the PAT for the USAID "very poor" line used to construct the PAT, however, differ in 7 of the 10 directly comparable countries. In some countries, the reason is unknown, but in others, the scorecard or the PAT make mistakes when implementing the poverty line. The discrepancies matter because potential targeting accuracy depends partly on the underlying poverty rate (Brandenburger and Furth, 2009). For the comparison here, the scorecard's lines are fixed (when they are in error) or adjusted so as to match the PAT's reported poverty rate (when the PAT is in error or when the reason for the discrepancy is unknown).[23] The appendix documents the fixes and adjustments.

[22] On $1.25/day, see Ravallion, Chen, and Sangraula (2009) and Sillers (2006). The median line divides people (not households) below a country's national poverty line into two equal-sized groups (U.S. Congress, 2004). The person-level poverty rate for the median line is half of the person-level rate for the national line. The PAT incorrectly derives the median line based on households instead of people (see appendix).

[23] To leave the scorecard's indicators and points untouched, the changes apply only to a scorecard's validation sample. The changes tilt accuracy comparisons in favor of the PAT because the PAT's construction—unlike the scorecard's—is tuned to this line.

3.2 Bias of estimates of poverty rates

Bias is the average difference in repeated samples between estimates and true values.[24] For example, suppose an archer shoots 10 arrows. If 5 are wide-left by 2 inches, 1 hits the bulls-eye, and 4 are wide-right by 1 inch, then the bias is wide-left by 0.6 inches, because $[(5 \text{ x } -2) + (1 \text{ x } 0) + (4 \text{ x } 1)] \div (5 + 1 + 4) = -0.6$.

Bias is an average in repeated samples. An estimator is *unbiased* if its bias is zero. For example, suppose that half an archer's arrows miss 3-inches-left and half miss 3-inches-right. The archer never hits the bulls-eye, but she is still unbiased because her average distance from the bulls-eye is zero.

3.2.1 Concepts of bias

The scorecard and PAT are compared in terms of several concepts of bias. The concepts correspond to the conditions in which the accuracy results can be extrapolated.

Testing accuracy means comparing estimates to true values. Short of doing new expenditure surveys, the only known true values are those from the national survey. Extrapolating measures of accuracy beyond this survey requires three assumptions.

The first is that the construction sample mirrors the population of households. Due to sampling variation[25] in the construction sample and in the national survey, this does not hold, although it matters less as sample size increases.

---

[24] Given N samples indexed by i, estimates e_i, and true values v_i, bias $= \sum_{i=1}^{N} \frac{e_i - v_i}{N}$.

[25] *Sampling variation* is what makes a given single sample—even if representative—an imperfect mirror of its population. Due to "luck of the draw", some sub-groups in the

The second assumption is that a tool is applied to groups that are representative of the same population that was used to construct the tool. In practice, this does not hold. The characteristics of a country's population change over time, so a nationally representative sample in one year is no longer nationally representative in a later year. This second assumption also does not hold when a tool is applied to participants of a local, pro-poor organization; if it did, the organization would not be local or pro-poor.

The third assumption is that the relationships between indicators and poverty do not change over time. This also does not hold in practice. For example, the cost-of-ownership for cell phones has fallen over time, so the relationship between poverty and cell-phone ownership is not the same now as it was five or ten years ago.

3.2.2 Ideal bias

Ideal bias is bias when all three assumptions hold. Although the assumptions do not hold in practice, ideal bias is still relevant because it can be measured and because lower ideal bias probably implies lower real-world bias. When this paper speaks of "bias" without other modifiers, it is speaking about ideal bias.

In the ideal case, both the scorecard and the PAT are unbiased. This follows from their structure, and it is most straightforwardly seen when the tools are applied

---

sample are under- or over-represented, even though the differences would average out in repeated samples. Like an estimator's bias, a sample's representativeness is defined in terms of averages in repeated samples. In a single sample, an unbiased estimator can miss its mark, and the sample can—and generally does—fail to represent its population.

in-sample, that is, when the validation sample is the same as the construction sample. In-sample, the three assumptions above hold.

The PAT's ideal bias is zero because the quantile q is selected to make in-sample bias zero. The scorecard's ideal bias is zero because poverty likelihoods are defined for a given score as the share of households in the construction sample who are poor.

3.2.3 Out-of-sample/in-time bias

Out-of-sample/in-time bias relaxes the first assumption but keeps the other two. While the construction sample is representative of a country's population of households, it is still a single sample that—thanks to sampling variation—does not mirror the population perfectly. The resulting bias can be seen by applying the scorecard *out-of-sample*, that is, when the validation sample differs from the construction sample. The second assumption holds because the validation sample is representative of the country's population. The test is *in-time* because the third assumption holds; the validation sample was collected at the same time as the construction sample, so the passage of time has not changed the relationships between indicators and poverty.

Out-of-sample/in-time bias is due to overfitting. A tool is *overfit* if it is tailored so closely to the construction sample that it captures not only some real patterns but

also some random patterns that, due to sampling variation, show up only in the construction sample but not in the validation sample or in the country's population.[26]

The scorecard reports this concept of bias for 52 of 53 countries, and the PAT reports it for 17 of 36 countries (Table 1). For the USAID "very poor" line, the average bias across countries is –0.1 percentage points for the scorecard and –0.3 percentage points for the PAT.[27]

When out-of-sample/in-time bias is known, it can be nullified; a user can get an unbiased estimate of a group's poverty rate by subtracting the known bias from the average poverty likelihood. Just as an archer who tends to miss by an average of 4 inches to the right will compensate by aiming 4 inches to the left of the bulls-eye, a user whose tool has a known bias of b percentage points can get an unbiased estimate by subtracting b from a group's average poverty likelihood.[28]

Looking at only the 10 countries where both tools use the same data and report out-of-sample/in-time results (Table 3), the scorecard's average bias is –0.5 percentage points, and the PAT's is +0.3.[29] The difference in the tools' bias is "statistically

[26] A tool may also be called *overfit* if it loses accuracy when it is applied to non-nationally representative samples (the second assumption fails) or when time changes the relationships between indicators and poverty (the third assumption fails).
[27] For both tools, the average *absolute* bias is about 1.1 percentage points.
[28] For example, if known bias is –1.1 percentage points and if the average poverty likelihood is 22.2 percent, then an unbiased estimate is 22.2 – (–1.1) = 23.3 percent. Bias is negative because the original estimate tends to be too low, so the adjustment increases the estimate. If bias is positive, then the adjustment decreases the estimate.
[29] The scorecard's average *absolute* bias is 1.1 percentage points, and the PAT's is 0.7.

significant" at the 90-percent level in 4 of the 10 countries,[30] and the PAT has smaller absolute bias in 7 of the 10 countries.

In sum, both tools have low out-of-sample/in-time bias, and they can be made unbiased by subtracting known bias from a group's average poverty likelihood.

3.2.4 Out-of-sample/out-of-time bias

In practice, the scorecard and the PAT are applied after the data was collected that was used for their construction. Even if a tool is applied to a nationally representative sample, the first and second assumptions do not hold because the population being represented changes with time, and the third assumption does not hold because the relationships between indicators and poverty also change with time.

In this *out-of-sample/out-of-time* case, both the scorecard and PAT are biased. While this bias can be nullified if it is known, this does not help much in practice. In particular, measuring this bias requires applying an existing tool to data from a new national expenditure survey. But because the characteristics of a country's population—and the indicator/poverty relationships—continue to change, the best way to reduce bias from now on is to construct a new tool with the new data. Of course, the new tool's out-of-sample/out-of-time bias is unknown and may not be like that of the old tool.

[30] Here, *statistically significant* means that the absolute value of the bias of at least one of the tools lies outside the 90-percent confidence interval of the absolute value of the bias of the other tool.

A tool can never catch up with out-of-sample/out-of-time bias. All that can be done is to construct a tool with the most-recent data, use indicators and approaches that are resistent to overfitting, and hope that the remaining (unknown) bias is not so large as to have a material effect on decisions that are informed by the estimates.

For the PAT, out-of-sample/out-of-time bias has not been measured. For the scorecard, it has been measured for 11 countries. This type of bias is not discussed further here because it cannot be compared between the scorecard and the PAT.

3.2.5 Out-of-group bias

Dropping the second assumption and keeping the first and third leads to *out-of-group bias*. This matters because the scorecard and PAT are constructed from nationally representative data but are applied to non-nationally representative sub-groups. In a sub-group, the relationships between indicators and poverty generally differ from those in the construction data.[31]

Common sub-groups are urban/rural, agricultural/non-agricultural, and sub-national regions. Of course, the participants of a pro-poor organization are a non-nationally representative sub-group, as they are both self-selected (as when a potential borrower chooses to apply for a microloan) and program-selected (as when a microlender approves loans for some potential borrowers but rejects others).

[31] In the context of poverty mapping, Tarozzi (2008) and Tarozzi and Deaton (2007) argue that sub-group differences can lead to large biases. Their point is parried by Demombynes, Elbers, and Lanjouw (2008) and Elbers, Lanjouw, and Leite (2008).

Out-of-group bias has not been measured for the scorecard nor the PAT. It is left as topic for future research.

3.2.6 Bias: summary

In ideal and simple cases, both the scorecard and the PAT are unbiased or can be made so. The two tools' bias is not known in more complex and realistic cases. Still, being unbiased in the simple, unrealistic cases is better than being biased in terms of what it suggests for possible accuracy in more complex, realistic cases.

3.3 Standard errors of estimates of poverty rates

The *standard error* σ measures the dispersion of estimates around their average value. In terms of the example archers above, the first is biased but has less dispersion (all arrows are within 1.6 inches of their average distance from the bull's-eye), while the second is unbiased but has more dispersion (all arrows miss the bull's-eye by 3 inches). Smaller standard errors are preferred to larger ones.

What are the standard errors of estimated poverty rates for the scorecard and PAT? Let the estimated poverty rate $\hat{p}_d$ of N_d households from a large national expenditure survey be their average directly *measured*—not indirectly *estimated*—poverty likelihood (100 percent if poor, 0 percent if non-poor). Then the standard error for an estimate of a poverty rate under direct measurement σ_d is $\sqrt{\frac{\hat{p}_d \cdot (1 - \hat{p}_d)}{N_d}}$.

The standard error σ_b for indirect measurement depends on the tool, country, and poverty line, so it is derived with bootstrapping rather than alegbra. Suppose N_b households indexed by i are bootstrapped from a validation sample. Let their estimated poverty likelihoods from a scorecard or PAT be e_i, and let their true, directly measured poverty likelihoods be v_i. Then $\sigma_b = \sum_{i=1}^{N_b} \sqrt{\frac{(e_i - v_i)^2}{N_b}}$.

Given a tool, country, and poverty line, the bootstrap standard error σ_b is (to within sampling variation) a multiple α of the direct-measurement standard error σ_d (Schreiner, 2008c). That is, $\sigma_b = \alpha \cdot \sigma_d$. If $\alpha < 1$, then indirect measurement via a scorecard or PAT has smaller standard errors than direct measurement via a survey.

α is useful because it lets users find desired sample sizes (before applying a tool) or report confidence intervals for estimated poverty rates (after applying a tool) using the direct-measurement formulas with σ_d replaced by $\alpha \cdot \sigma_d$. Also, α is a concise measure of relative standard error, facilitating comparisons here.

Table 3 reports α for the scorecard and PAT in the 10 countries where both use the same national-survey data and both report bootstrapped out-of-sample/in-time tests. The scorecard reports the average α across 1,000 bootstraps of $N_b = 2^8, 2^9, \ldots, 2^{14}$. The PAT's α is $\frac{|\, c_+ - c_- \,|}{2} \cdot \left(1.96 \sqrt{\frac{p_v \cdot (1 - p_v)}{N_v}} \right)^{-1}$, where c_- and c_+ are the reported bootstrapped endpoints of the 95-percent confidence interval of the PAT estimate, 1.96 is the z-value for 95-percent confidence intervals, and p_v is the true poverty rate (and N_v the number of households) in the PAT's validation sample.

The scorecard has smaller standard errors in 8 of 10 countries. Its average α of 0.91 is about 20 percent less than the PAT's average α of 1.12.

The result is similar across all countries with out-of-sample/in-time tests (Table 1); α averages 0.95 in the 52 scorecard countries and 1.18 in the 17 PAT countries.

3.4 Targeting accuracy

Targeting accuracy is measured as *inclusion* (the share of all households who are truly poor and correctly classified as poor) and *exclusion* (the share of all households who are truly non-poor and correctly classified as non-poor). The sum of inclusion and exclusion is the *hit rate*. Higher inclusion and higher exclusion are better than lower.

Targeting accuracy can also be seen in terms of *undercoverage* (the share of all households who are truly poor but mistakenly classified as non-poor) and *leakage* (the share of all households who are truly non-poor but mistakenly classified as poor, Figure 4). It is easier, however, to think in terms of successes (inclusion and exclusion) than mistakes (undercoverage and leakage).[32]

A tool's potential targeting accuracy depends partly on a population's poverty rate; higher hit rates are easier to achieve as poverty rates are further from 50 percent. Thus, comparisons are cleanest in the 10 countries in which the scorecard and PAT use

[32] When setting a targeting cut-off, an organization will generally consider all four possible targeting outcomes, weighting each according to its net benefit. In this paper, targeting accuracy is compared across the scorecard and PAT while holding constant both the share of all households who are targeted and the underlying poverty rate. In this case, it is sufficient to look at the hit rate.

the same data, have out-of-sample/in-time tests, and have the same poverty rates for the USAID "very poor" line. The scorecards' cut-offs are set so as to target the same share of households in its validation sample as the PAT does.[33]

On average in the 10 countries in Table 3, the PAT has better inclusion by 1.2 percentage points, and the scorecard has better exclusion by 0.5 percentage points. Combining these differences for inclusion and exclusion, the PAT has a better hit rate by 0.7 percentage points.

Head-to-head, the PAT has better inclusion in 8 of 10 countries (five of these differences—all in favor of the PAT—are "statistically significant" in that the scorecard's inclusion lies outside the PAT's 90-percent confidence interval). The PAT has better exclusion in 5 of 10 countries (one of these differences is "statistically significant".) For the hit rate, the PAT is better in 8 of 10 countries (none of the differences is "statistically significant").

[33] The scorecard has 20 possible cut-offs, so targeting the same share of households as the PAT requires interpolating between cut-offs. In the case of Peru, the PAT targets 18.9 percent of households in its validation sample (that is, 18.9 percent have 2.71828^{score} below the median line). Inclusion is 10.7 percent, and exclusion is 76.3. For the scorecard, a cut-off score of 29 or less targets 14.3 percent of households, with inclusion of 9.0 and exclusion of 78.8. The next-higher cut-off of 34 or less targets 22.2 percent of households, with inclusion of 12.1 and exclusion of 74.0. To match the PAT's targeting of 18.9 percent of households, the scorecard targets all households scoring 29 or less and a random (18.9 – 14.3) ÷ (22.2 – 14.3) = 58.2 percent of households scoring 30–34. This gives inclusion of 10.8 percent and exclusion of 76.0.

In sum, the PAT correctly classifies about one more household per 100. This targeting edge probably follows from the PAT's use of quantile regression and (mostly) from its use of 15–18 indicators rather than 10.

The PAT's developers say that the PAT is too inaccurate to target individual households. FHI360 (2013, p. 9) warns, "Do not make decisions that affect household-member lives based solely on PAT data (that is, do not use PAT-calculated consumption information as the only criteria for household targeting)." If this warning is appropriate, then it would also apply to the scorecard.[34]

In counterpoint, Schreiner (2008a) suggests that household-level targeting is a legitimate potential use of scorecards. Whether a tool is accurate enough for a specific purpose in a specific context depends on the benefits of successful inclusion and exclusion and on the costs of mistaken undercoverage and leakage, as well as on the alternatives for accomplishing the same goals. Several of the world's largest lower-income countries already use scorecard- and PAT-like tools—called "proxy means tests"—to target social transfers to the poor.[35] A blanket statement that PATs should never be used for targeting seems as unlikely to be true as would a claim that

[34] The developers of another poverty-measurement tool (poverty maps) also say that it is too inaccurate to target individual households (Elbers, Lanjouw, and Lanjouw, 2003; Demombynes *et al.*, 2004), although Elbers *et al.* (2007) seems to back off a bit.
[35] World Bank (2012, Indonesia); Fernandez (2012, Philippines); Camacho and Conover (2011, Colombia); Sharif (2009, Bangladesh); World Bank (2009, Pakistan); Mostafa and da Silva (2007, Brazil); and Coady (2006, Mexico).

scorecards should always be used for targeting. The key is that targeting accuracy be transparent so that users can see the pros and cons and make informed choices.

3.5 BPAC = Inclusion

This paper discusses accuracy as bias, standard error, inclusion, and exclusion. These concepts are standard in statistics and in the scoring literature. USAID certifies PATs, however, based on the Balanced Poverty Accuracy Criterion. How do the scorecard and PAT differ in terms of BPAC?

IRIS Center (2005) defines BPAC as inclusion minus the absolute value of bias.[36] If the targeting cut-off is set to be the same as the poverty line—as it is in the PAT documentation—then bias is the same as undercoverage minus leakage:

$$\begin{aligned} \text{BPAC} &= \text{Inclusion} - |\ \text{Bias}\ | \\ &= \text{Inclusion} - |\ \text{Undercoverage} - \text{Leakage}\ |. \end{aligned}$$

But the PAT's bias is (Undercoverage – Leakage) only if the same cut-off is used both for targeting and for estimating a household's poverty likelihood (100 percent if 2.71828^{score} is below the cut-off, 0 percent otherwise). The PAT documentation uses the USAID "very poor" poverty line for both cut-offs, but the cut-offs could differ. For example, suppose that 30 percent of households in a country are below the USAID "very poor" line. Nothing dictates that a pro-poor organization—if it targets at all—

[36] BPAC typically multiplies (Inclusion– | Bias |) by [100 ÷ (Inclusion + Undercoverage)]. If two tools are applied to households with the same poverty rate—as in Table 3—then the second term can be dropped without affecting the comparison.

must target the lowest-scoring 30 percent of households. It might want to target the "ultra-poor" (say, the lowest-scoring 10 percent) or the lowest-scoring 50 percent.

When a PAT uses different cut-offs for targeting and estimating poverty likelihoods, bias is no longer equal to the difference between undercoverage and leakage. Thus, BPAC is better expressed as Inclusion – | Bias |.[37]

This makes it simple to compare BPAC across the scorecard and PAT. Both tools are unbiased in the ideal case, and both can be made to be unbiased in the out-of-sample/in-time case. With no bias, BPAC = inclusion. It is more transparent to talk about inclusion directly, rather than cloaking it in BPAC. For the 10 countries in Table 3, the PAT's inclusion (and thus its BPAC) exceeds the scorecard's on average by 1.2 percentage points.

[37] The poverty scorecard does not use cut-offs to estimate poverty likelihoods. Thus, any cut-off for targeting always differs from the (non-existent) cut-off for estimating poverty likelihoods, and bias is not (Undercoverage – Leakage).

4. Conclusion

The scorecard and the PAT are simple, low-cost ways to measure poverty. They have similar accuracy.[38] The scorecard is more widely available, more recent, and more transparent.

Does take-up differ? Each tool has a support organization (Grameen Foundation for the PPI®-branded scorecard, and USAID for the PAT) that helps users with training, documentation, and other aids.[39] There are no complete counts of active users, however, because the tools can be downloaded at no cost and without registering and because some former users are no longer active.

As of 27 February 2014, there were about 140 verified, active users of the PPI®-branded scorecard.[40] Among the 405 microfinance organizations who voluntarily reported social-performance information to the MIX Market in 2009–10 (Pistelli, Simanowitz, and Thiel, 2011, p. 17), 58 used the PPI®, and 12 used the PAT.

By law, 50 percent of USAID's microenterprise funding must benefit the "very poor". As an input to its annual reporting to Congress, USAID requires its microenterprise partners to use the PAT to estimate the share of their participants who

[38] Before there was any evidence of the kind presented here, Schreiner (2005) anticipated this result based on the "flat maximum" phenomenon in which reasonable index-based tools have similar accuracy (Hand, 2006; Baesens *et al.*, 2003; Lovie and Lovie, 1986; Kolesar and Showers, 1985; Stillwell, Barron, and Edwards, 1983; Dawes, 1979; and Wainer, 1976).

[39] For take-up, support is important. Comparing the support available for the two tools is beyond the scope of this paper, as is comparing their overall financial costs.

[40] Grameen Foundation, `progressoutofpoverty.org/ppi-users`.

are "very poor".[41] Despite this mandate, the report for 2011 says "up-take by USAID's partners has been relatively low". Table 4 shows that from 2007 to 2012, the number of countries covered by a PAT increased from 17 to 38, but the number of partners who reported PAT results decreased from 31 to 4.[42]

In the same stretch, participants' estimated poverty rate increased from 22 to 56 percent. Does this mean that the law that led to the development of the PAT fulfilled its basic aim, that is, to prompt USAID to take new partners with greater poverty focus and to pressure existing partners to serve the "very poor" disproportionately more?

In its report for 2012 (p. 6), USAID says that the 56-percent poverty-rate estimate "may be too limited to be meaningful" because the four PAT-reporters are only five percent of all reporting partners. Likewise, USAID (p. 8) says that 2011's estimate of 38 percent cannot be taken at face-value because "year-to-year changes in the overall percentage of 'very poor' clients . . . are largely driven by the changing mix of partners that report results from the PAT."

In short, so few partners report PAT results that USAID does not know whether it is making progress towards its legal mandate. USAID lays the blame on the law's "very poor" poverty line. Annex C in the report for 2007 says that this line is

[41] In absence of a waiver, this includes projects and partner organizations receiving at least $100,000 from USAID in a fiscal year for microenterprise in countries with a PAT.
[42] The annual reports do not say how many partners should have reported but did not. Before 2012, all partners were sent questionnaires, but 2012 covered only a sample. Reporting for 2010 fell in part due to a temporary suspension in the requirement to use the PAT due to a legal challenge under the Paperwork Reduction Act.

"unrealistically stringent" and therefore "USAID could not meet the legislative [mandate] . . . through *any* [italics original] reallocation of funds among its current partners" and that attempting to deepen poverty outreach by taking on new, higher-cost partners with greater poverty outreach would harm partners' sustainability and harm the broad-based economic growth that is also part of USAID's mandate. The report for 2008 (p. 14) says "Two years of results using PATs strongly suggests that the [mandate] *cannot* [italics original] be reached without inflicting undesirable side effects on sustainability and economic development. In short, USAID sees no realistic prospect of reaching the [legal mandate]."

USAID makes a case that the law's standard is too strict. But the law is the law. Like a driver pulled-over for speeding who tells the ticketing officer that the speed limit should be higher, USAID has the right to protest that the "very poor" line should be higher. But even if USAID is right, it still has to obey the law.

USAID's annual reports do not describe doing anything specific—beyond "working hard", mandating poverty measurement, and creating PATs and supporting their use—to increase the share of partners' participants who are "very poor". The decrease in the number of PAT-reporters suggests that partners do not expect consequences for not reporting. And with so few partners reporting, there is no way to know whether USAID is picking new partners with greater depth of outreach than existing partners or whether USAID is pressuring existing partners to work to find new

ways to increase depth of outreach over time while still being sustainable and promoting broad-based economic growth.

The law sets a high bar, and it takes work to do new things and to take risks to get better in one aspect (poverty outreach) and not get worse in other aspects (sustainability and economic growth). But you do not get better if you do not try. The issue is not the PAT nor the "very poor" poverty line; it is the lack of evidence that USAID is trying to make progress towards the fulfillment of its legal mandate.

References

Baesens, Bart; Van Gestel, Tony; Viaene, Stijn; Stepanova, Maria; Suykens, Johan A.K.; and Jan Vanthienen. (2003) "Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring", *Journal of the Operational Research Society*, Vol. 54, pp. 627–635.

Brandenburger, Thomas; and Alfred Furth. (2009) "Cumulative Gains Model Quality Metric", *Journal of Applied Mathematics and Decision Sciences*, doi: 10.1155/2009/868215.

Caire, Dean; Schreiner, Mark; Chen, Shiyuan; and Gary Woller. (2009) "A Simple Poverty Scorecard for Nepal", `microfinance.com/English/Papers/Scoring_Poverty_Nepal_EN_2003.pdf`, retrieved 17 January 2014.

Camacho, Adriana; and Emily Conover. (2011) "Manipulation of Social-Program Eligibility", *American Economic Journal: Economic Policy*, Vol. 3, No. 2, pp. 41–65.

Central Bureau of Statistics. (2005) "Poverty Trends in Nepal: (1995–96 and 2003–04)", National Planning Commission Secretariat, Government of Nepal, `siteresources.worldbank.org/INTLSMS/Resources/3358986-1181743055198/3877319-1181925143929/NepalPovertyTrendsforPovertyAssessment.pdf`, retrieved 17 January 2014.

Centro de Estudios Distributivos, Laborales y Sociales; and World Bank. (2012) "A Guide to the SEDLAC Socio-Economic Database for Latin America and the Caribbean", Buenos Aires: Universidad Nacional de la Plata, `sedlac.econo.unlp.edu.ar/download.php?file=archivos_upload_items_metodologia/Guide_14_english.pdf`, retrieved 17 January 2014.

Coady, David. (2006) "The Welfare Returns to Finer Targeting: The Case of the PROGRESA Program in Mexico", *International Tax and Public Finance*, Vol. 13, pp. 217–239.

CONEVAL. (2009) "Aplicación de la Metodología para la Medición de la Pobreza por Ingresos y Pruebas de Hipótesis 2008", Mexico, D.F.: Consejo Nacional de Evaluación de la Política de Desarrollo Social, `coneval.gob.mx/rw/resource/coneval/med_pobreza/3495.pdf`, retrieved 17 January 2014.

Dawes, Robyn M. (1979) "The Robust Beauty of Improper Linear Models in Decision-Making", *American Psychologist*, Vol. 34, No. 7, pp. 571–582.

Demombynes, Gabriel; Elbers, Chris; and Peter Lanjouw. (2008) "How Good a Map? Putting Small-Area Estimation to the Test", *Rivista Internazionale di Scienze Sociali*, Vol. 116, No. 4, pp. 465–494.

_____; Elbers, Chris; Lanjouw, Jenny; Lanjouw, Peter; Mistiaen, Johan; and Berk Özler. (2004) "Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries", pp. 154–176 in Anthony Shorrocks and Rolph van der Hoeven (eds.) *Growth, Inequality, and Poverty*, Oxford University Press.

Dirección General de Estadística y Censos. (2009) "Encuesta de Hogares de Propósitos Múltiples, 2008", Delgado.

Drèze, Jean; and Reetika Khera. (2010) "The BPL Census and a Possible Alternative", *Economic and Political Weekly*, Vol. 45, No. 9, pp. 54–63.

Elbers, Chris; Fujii, Tomoki; Lanjouw, Peter; Özler, Berk; and Wesley Yin. (2007) "Poverty Alleviation through Geographic Targeting: How Much Does Disaggregation Help?", *Journal of Development Economics*, Vol. 83, pp. 198–213.

_____; Lanjouw, Jean O.; and Peter Lanjouw. (2003) "Micro-Level Estimation of Poverty and Inequality", *Econometrica*, Vol. 71, No. 1, pp. 355–364.

_____; Lanjouw, Peter; and Phillippe George Leite. (2008) "Brazil within Brazil: Testing the Poverty-Map Methodology in Minas Gerais", World Bank Policy Research Working Paper No. 4513, Washington, D.C., `elibrary.worldbank.org/doi/book/10.1596/1813-9450-4513`, retrieved 17 January 2014.

Fernandez, Luisa. (2012) "Design and Implementation Features of the National Household-Targeting System in the Philippines", World Bank Philippine Social Protection Note. No. 5, `ideas.repec.org/p/wbk/hdnspu/70149.html`, retrieved 17 January 2014.

FHI360. (2013) "USAID Poverty Assessment Tool (PAT) Data-Analysis Guide", `povertytools.org/USAID_documents/Manual/PATDataAnalysisManual.pdf`, retrieved 17 January 2014.

Ghana Statistical Service. (2007) "Patterns and Trends in Poverty in Ghana, 1991–2006", Accra, `catalog.ihsn.org/index.php/catalog/63/download/28521`, retrieved 17 January 2014.

Hand, David J. (2006) "Classifier Technology and the Illusion of Progress", *Statistical Science*, Vol. 22, No. 1, pp. 1–15.

Institut National de la Statistique. (2006) "Preliminary Poverty Update Report", Kigali, 196.44.242.24/eicv/survey0/data/docs/studies/Master%20Report.pdf, retrieved 17 January 2014.

Instituto Nacional de Estadística y Censos. (2007) "Homologación Metodológica de Cálculo de Pobreza: Desigualdad e Indicadores Sociales, a Partir de la Encuesta de Condiciones de Vida (ECV)", Quito, inec.gov.ec/c/document_library/get_file?folderId=80249&name=DLFE-3325.pdf, retrieved 19 July 2011.

Instituto Nacional de Estadística e Informática. (2011) "Evolución de la Pobreza al 2010", Lima, dger.minem.gob.pe/ArchivosNormasTecnicas/Indicadores%20Sociales%20Ind_Pobreza%202010.pdf, retrieved 17 January 2014.

Instituto Nacional de Información de Desarrollo (2007) "Perfil y Características de los Pobres en Nicaragua 2005", Managua, inide.gob.ni/bibliovirtual/publicacion/Informe%20EMNV%202009.pdf, retrieved 17 January 2014.

IRIS Center. (2011a) "Poverty Assessment Tool Submission for Nicaragua", 16 September, povertytools.org/countries/Nicaragua/USAID_PAT_Nicaragua.pdf, retrieved 17 January 2014.

_____. (2011b) "Poverty Assessment Tool Submission for Nigeria", 5 August, povertytools.org/countries/Nigeria/USAID_PAT_Nigeria.pdf, retrieved 17 January 2014.

_____. (2011c) "Poverty Assessment Tool Submission for Peru", 15 September, povertytools.org/countries/Peru/USAID_PAT_Peru_certification_15Sept2011.pdf, retrieved 17 January 2014.

_____. (2011d) "Poverty Assessment Tool Submission for Rwanda", 30 June, povertytools.org/countries/Rwanda/USAID_PAT_Rwanda.pdf, retrieved 17 January 2014.

_____. (2011e) "Poverty Assessment Tool Submission for Tanzania", 16 September, povertytools.org/countries/Tanzania/USAID_PAT_TZ.pdf, retrieved 17 January 2014.

_____. (2010a) "Poverty Assessment Tool Submission for Ecuador", 30 June, povertytools.org/countries/Ecuador/USAID_PAT_Ecuador.pdf, retrieved 17 January 2014.

_____. (2010b) "Poverty Assessment Tool Submission for El Salvador", 18 June, povertytools.org/countries/El_Salvador/USAID_PAT_El_Salvador.pdf, retrieved 17 January 2014.

_____. (2010c) "Poverty Assessment Tool Submission for Ghana", 4 June, povertytools.org/countries/Ghana/USAID_PAT_Ghana.pdf, retrieved 17 Jaunary 2014.

_____. (2010d) "Poverty Assessment Tool Submission for Kenya", 20 October, povertytools.org/countries/Kenya/USAID_PAT_Kenya.pdf, retrieved 17 January 2014.

_____. (2010e) "Poverty Assessment Tool Submission for Mexico", July 19, povertytools.org/countries/Mexico/USAID_PAT_Mexico.pdf, retrieved 17 Jaunary 2014.

_____. (2009a) "Poverty Assessment Tool Submission for Cambodia", 30 September, povertytools.org/countries/Cambodia/USAID_PAT_Cambodia.pdf, retrieved 17 January 2014.

_____. (2009b) "Poverty Assessment Tool Submission for Ethiopia", 17 December, povertytools.org/countries/Ethiopia/USAID_PAT_Ethiopia.pdf, retrieved 17 January 2014.

_____. (2009c) "Poverty Assessment Tool Submission for Malawi", 30 September, povertytools.org/countries/Malawi/USAID_PAT_Malawi.pdf, retrieved 17 January 2014.

_____. (2009d) "Poverty Assessment Tool Submission for Nepal", 1 June, povertytools.org/countries/Nepal/USAID_PAT_Nepal.pdf, retrieved 17 January 2014.

_____. (2006) "Poverty Assessment Tool Submission for Guatemala", 31 October, povertytools.org/countries/Guatemala/USAID_PAT_Guatemala2013.pdf, retrieved 17 January 2014.

_____. (2005) "Notes on Assessment and Improvement of Tool Accuracy", povertytools.org/other_documents/AssessingImproving_Accuracy.pdf, retrieved 17 January 2014.

Knowles, James C. (2006) "A New Set of Poverty Estimates for Cambodia, 1993/94 to 2004", Revision 28.11.06, Report to the EAS Country Units of the World Bank, jkinusa@aol.com.

Kolesar, Peter; and Janet L. Showers. (1985) "A Robust Credit-Screening Model Using Categorical Data", *Management Science*, Vol. 31, No. 2, pp. 124–133.

Lovie, Alexander D.; and Patricia Lovie. (1986) "The Flat-Maximum Effect and Linear Scoring Models for Prediction", *Journal of Forecasting*, Vol. 5, pp. 159–168.

Matul, Michal; and Sean Kline. (2003) "Scoring Change: Prizma's Approach to Assessing Poverty", Spotlight Note No. 4, Warsaw: Microfinance Centre for Central and Eastern Europe and the New Independent States, impact.org/sites/default/files/mfc_sn4.pdf, retrieved 17 January 2014.

Ministry of Planning. (2006) *A Poverty Profile of Cambodia, 2004*, Phnom Penh: Royal Government of Cambodia, mop.gov.kh/Others/PovertyProfile/tabid/191/Default.aspx, retrieved 17 January 2014.

Mostafa, Joana; and Kátia Cristina da Silva. (2007) "Brazil's Single Registry Experience: A Tool for Pro-Poor Social Policies", www.ipc-undp.org/doc_africa_brazil/Webpage/missao/Artigos/CadastroUnicoJoanaMostafa.pdf, retrieved 17 January 2014.

National Bureau of Statistics. (2009) *Main Report: Household Budget Survey 2007, Tanzania Mainland*, Dar es Salaam, nbs.go.tz/tnada/index.php/catalog/2/download/39, retrieved 17 January 2014.

National Bureau of Statistics. (2005) "Poverty Profile for Nigeria", Abuja.

National Institute of Statistics. (2007) "Basic Report on Well-Being in Kenya", Ministry of Planning and National Development, statistics.knbs.or.ke/nada/index.php/catalog/8/download/35, retrieved 17 January 2014.

Neihaus, Paul; Atanassova, Antonia; Bertrand, Marianne; and Sendhil Mullainathan. (2013) "Targeting with Agents", *Amercan Economic Journal: Economic Policy*, Vol. 5, No. 1, pp. 206–38.

Pistelli, Micol; Simanowitz, Anton; and Veronika Thiel. (2011) "A Survey of 405 MFIs Reporting to MIX in 2009/10", themix.org/sites/default/files/MBB-%20SoPinSPReporting%20and%20Management_FINAL.pdf, retrieved 17 January 2014.

Poon, Martha. (2007) "Scorecards as Devices for Consumer Credit: the Case of Fair, Isaac & Company Incorporated", *Sociological Review*, Vol. 55, Issue supplement s2, pp. 284–306.

Ravallion, Martin; Chen, Shaohua; and Prem Sangraula. (2009) "Dollar-a-Day Revisited", *World Bank Economic Review*, Vol. 23, No. 2, pp. 163–184.

SAS Institute Inc. (2004) "The LOGISTIC Procedure: Rank Correlation of Observed Responses and Predicted Probabilities", in *SAS/STAT User's Guide, Version 9*, Cary, NC., support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_logistic_sect035.htm, retrieved 17 January 2014.

Schreiner, Mark. (2013a) "A Simple Poverty Scorecard for Nepal", microfinance.com/English/Papers/Scoring_Poverty_Nepal_2010_EN.pdf, retrieved 17 January 2014.

_____. (2013b) "A Simple Poverty Scorecard for Nicaragua", microfinance.com/English/Papers/Scoring_Poverty_Nicaragua_2009_EN.pdf, retrieved 17 January 2014.

_____. (2012) "A Simple Poverty Scorecard for Peru", microfinance.com/English/Papers/Scoring_Poverty_Peru_2010_EN.pdf, retrieved 17 January 2014.

_____. (2011a) "A Simple Poverty Scorecard for Kenya", microfinance.com/English/Papers/Scoring_Poverty_Kenya_EN_2005.pdf, retrieved 17 January 2014.

_____. (2011b) "A Simple Poverty Scorecard for Malawi", microfinance.com/English/Papers/Scoring_Poverty_Malawi_2004_EN.pdf, retrieved 17 January 2014.

_____. (2011c) "A Simple Poverty Scorecard for Tanzania", microfinance.com/English/Papers/Scoring_Poverty_Tanzania_2007_EN.pdf, retrieved 17 January 2014.

_____. (2010a) "A Simple Poverty Scorecard for El Salvador", microfinance.com/English/Papers/Scoring_Poverty_El_Salvador_EN_2008.pdf, retrieved 17 January 2014.

_____. (2010b) "A Simple Poverty Scorecard for Rwanda", microfinance.com/Francais/Articles/Scoring_Poverty_Rwanda_FR_2005.pdf, retrieved 17 January 2014.

_____. (2010c) "How to Transform Coefficients from a Regression on Categorical Indicators to Non-Negative Integer Points Whose Sum Ranges from 0 to x", memo to Dean Caire, 10 June.

_____. (2009a) "A Simple Poverty Scorecard for Cambodia", microfinance.com/English/Papers/Scoring_Poverty_Cambodia_EN_2004.pdf, retrieved 17 January 2014.

_____. (2009b) "A Simple Poverty Scorecard for Mexico", microfinance.com/English/Papers/Scoring_Poverty_Mexico_2008_EN.pdf, retrieved 17 January 2014.

_____. (2008a) "A Simple Poverty Scorecard for Ecuador", microfinance.com/English/Papers/Scoring_Poverty_Ecuador_EN_2005.pdf, retrieved 17 January 2014.

_____. (2008b) "A Simple Poverty Scorecard for Nigeria", microfinance.com/English/Papers/Scoring_Poverty_Nigeria_EN_2003.pdf, retrieved 17 January 2014.

_____. (2008c) "A Simple Poverty Scorecard for Peru", microfinance.com/English/Papers/Scoring_Poverty_Peru_May_2008.pdf, retrieved 17 January 2014.

_____. (2005) "Response to IRIS questions on poverty scorecards", microfinance.com/English/Papers/Scoring_Poverty_Response_to_IRIS.pdf, retrieved 17 January 2014.

_____; and Shiyuan Chen. (2009) "A Simple Poverty Scorecard for Ethiopia", microfinance.com/English/Papers/Scoring_Poverty_Ethiopia.pdf, retrieved 17 January 2014.

_____; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Poverty Scorecards: Lessons from a Microlender in Bosnia-Herzegovina", microfinance.com/English/Papers/Scoring_Poverty_in_BiH_Short.pdf, retrieved 17 January 2014.

_____; and Gary Woller. (2010a) "A Simple Poverty Scorecard for Ghana", microfinance.com/English/Papers/Scoring_Poverty_Ghana_EN_2005.pdf, retrieved 17 January 2014.

_____; and Gary Woller. (2010b) "A Simple Poverty Scorecard for Guatemala", microfinance.com/English/Papers/Scoring_Poverty_Guatemala_EN_2006.pdf, retrieved 17 January 2014.

_____; and Gary Woller. (2010c) "A Simple Poverty Scorecard for Nicaragua", microfinance.com/English/Papers/Scoring_Poverty_Nicaragua_2005_EN.pdf, retrieved 17 January 2014.

Sharif, Iffath Anwar. (2009) "Building a Targeting System for Bangladesh Based on Proxy-Means Testing", World Bank Social Protection Discussion Paper No. 0914, Washington, D.C., siteresources.worldbank.org/SOCIALPROTECTION/Resources/SP-Discussion-papers/Safety-Nets-DP/0914.pdf, retrieved 17 January 2014.

Sillers, Don. (2006) "National and International Poverty Lines: An Overview", Washington, D.C.: USAID, pdf.usaid.gov/pdf_docs/Pnadh069.pdf, retrieved 17 January 2014.

Stillwell, William G.; Barron, F. Hutton; and Ward Edwards. (1983) "Evaluating Credit Applications: A Validation of Multi-Attribute Utility-Weight Elicitation Techniques", *Organizational Behavior and Human Performance*, Vol. 32, pp. 87–108.

Tarozzi, Alesandro. (2008) "Can Census Data Alone Signal Heterogeneity in the Estimation of Poverty Maps?", *Journal of Development Economics*, Vol. 95, No. 2, pp. 170–185.

_____; and Angus Deaton. (2007) "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", *Review of Economics and Statistics*, Vol. 91, No. 4, pp. 773–792.

USAID (various years) *Microenterprise Results Reporting*, with methodology and statistical annexes, fiscal years 2007 to 2012, mrreporting.org/publications/, retrieved 27 February 2014.

United States Congress. (2004) "Microenterprise Results and Accountability Act of 2004 (HR 3818 RDS)", November 20, smith4nj.com/laws/108-484.pdf, retrieved 17 January 2014.

Wainer, Howard. (1976) "Estimating Coefficients in Linear Models: It Don't Make No Nevermind", *Psychological Bulletin*, Vol. 83, pp. 223–227.

World Bank. (2012) *Targeting Poor and Vulnerable Households in Indonesia*, aid.dfat.gov.au/countries/eastasia/indonesia/Documents/prsp-report-targeting-poor-ipm.pdf, retrieved 17 January 2014.

_____. (2009) "Poverty Scorecard for Pakistan: A Recommended Approach for Targeting the Poor", Washington, D.C.

_____. (2008) "International Comparison Project: Tables of Results", Washington, D.C., siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, retrieved 17 January 2014.

_____. (2007) "Malawi Poverty and Vulnerability Assessment Investing in Our Future, Volume I: Synthesis of Main Findings and Recommendations" and "Volume II: Detailed Background Papers" and "Appendices", World Bank Report No. 36546-MW, openknowledge.worldbank.org/handle/10986/7557, retrieved 17 January 2014.

_____. (2006) *Cambodia: Halving Poverty By 2015? Poverty Assessment, 2006*, Washington, D.C., go.worldbank.org/LWTE2DIAI1, retrieved 17 January 2014.

_____. (2003) *Poverty in Guatemala*, Report No. 24221–GU, Washington, D.C., go.worldbank.org/RKDSQ3FZZ0, retrieved 17 January 2014.

Zeller, Manfred. (2004) "Review of Poverty Assessment Tools", pdf.usaid.gov/pdf_docs/PNADH120.pdf, retrieved 17 January 2014.

_____; Alcaraz V., Gabriela; and Julia Johannsen. (2005) "Developing and Testing Poverty-Assessment Tools: Results from Accuracy Tests in Peru", povertytools.org/other_documents/Peru%20Accuracy%20Report.pdf, retrieved 17 January 2014.

Appendix:
Errors in Implementing Poverty Lines

The scorecard and PAT implement some poverty lines incorrectly or err in other ways. For the USAID "very poor" line, the scorecard has errors in four of 13 countries, and the PAT has errors in 14 of 15 countries (Table A1).[43] This is disconcerting. Furthermore, the scorecard's estimates of $1.25/day poverty usually differ from those of the World Bank's PovcalNet. In these cases, however, it appears that the scorecard's corrected estimates are to be preferred.

The scorecard's errors are:

- Not following a country's regional poverty lines (Ecuador)
- Using the wrong time units for a poverty line (Ethiopia, Malawi)
- Not adjusting for regional-price deflators that do not average exactly 1.0 (Kenya, Malawi, Nicaragua, Tanzania)
- Dropping 20 households who reported not owning any assets (Kenya)
- Using the wrong household-roster data (Tanzania)

The PAT's errors are:

- Deriving the median line terms of households instead of people (all countries)
- Not following a country's regional poverty lines (Ghana, Malawi, Nepal, Peru)
- Using the wrong time units for a poverty line (Ethiopia, Nepal, Nigeria, Rwanda)
- Using the wrong data sub-set (Cambodia)
- Using the wrong household-roster data (Tanzania)

[43] This appendix discusses the 15 countries in which the two tools are based on the same survey data. The head-to-head accuracy comparisons in the main text and Table 3 discuss only 10 of the 15. Two (Cambodia and Tanzania) are not comparable head-to-head because of non-replicable errors by the PAT. Three others (Ethiopia, Guatemala, and Nepal) are not comparable head-to-head because the PAT does not report out-of-sample/in-time accuracy. These five are included in this appendix because their implementaton of poverty lines and their data handling can be compared.

The errors matter for three reasons. First, comparing the scorecard and the PAT head-to-head in a given country requires that both tools be applied to households with the same poverty rate. For this paper, this is accomplished by fixing the scorecard's poverty lines (when they are in error) or adjusting them to match the PAT's poverty rate (when the PAT is in error, or when a discrepancy is unexplained).

Second, existing poverty-rate estimates based on the mistaken lines are inaccurate.

Third, the frequency of errors in the lines and countries checked here suggests that other errors lurk in other lines and other countries which have not been checked. This casts doubt on existing poverty-rate estimates in general.

Not all is lost. The implementation of national lines can be checked by comparing a country's official person-level poverty rate against that found by a tool when applied to data from a national expenditure survey.[44] The scorecard implements the national line correctly in 13 of 14 countries,[45] while the PAT implements it correctly in 8 of 14 of countries (Table A2).[46]

[44] The definition of a country's national line may have flaws, but it is the most relevant line in that country and the only solid benchmark for checking correct implementation. This paper assumes that correct median lines and correct $1.25/day lines should mimic how the national line is adjusted for price differences across poverty-line regions.

[45] Ethiopia does not have a national poverty line. The scorecard's one error is to adjust Ecuador's line for regional food-price differences even though the official line does not.

[46] The PAT usually does not report a person-level rate for the national line. But if the scorecard's person-level rate for the national line matches the official rate, then the PAT's implementation can be inferred to be correct if its household-level rate for the national line (found as twice its household-level rate for the median line) matches the scorecard's household-level rate.

The median line is also easily checked; its person-level poverty rate should be half that of the national line. This is the case in all countries for which the scorecard's national line is correct. The PAT derives the median line incorrectly in all countries.

For lack of established criteria, benchmarks, and documentation, it is difficult to check the correctness of implementation of the $1.25/day line. Only the World Bank's PovcalNet[47] has $1.25/day estimates for more countries (129) than the scorecard (53). Although PovcalNet is the main source of estimates of worldwide progress toward the Millenium Development Goal of halving the number of people under $1.25/day, its implementation of this line often seems to be in error for the 15 countries here (Table A3). In general, the scorecard's $1.25/day estimates are to be preferred because they are more transparent and more likely to be correct.[48] Unlike PovcalNet, the scorecard does the following in all of its countries:

- Uses household-level data from a national expenditure survey
- Reports the $1.25/day line in local currency as of a specific time period
- Replicates a country's official person-level poverty rate for the national line[49]
- Adjusts for price differences across poverty-line regions defined by a country
- Adjusts 2005 PPP factors over time to match the time-unit of expenditure
- Documents the above elements of the derivation of the $1.25/day line

---

[47] "PovcalNet: The On-Line Tool for Poverty Measurement Developed by the Development Research Group of the World Bank", `iresearch.worldbank.org/PovcalNet/index.htm`, retrieved 15 January 2014.

[48] Table A3 and the country-specific notes in this appendix. These notes have been shared with PovcalNet.

[49] PovcalNet may sometimes use different definitions or different survey data than that used by a country's government. While these differences might make sense, PovcalNet's documentation is often incomplete, and a reasonable default is to follow what a country's government does.

The scorecard's $1.25/day estimates are not always better than PovcalNet's, and the scorecard covers fewer countries for fewer years.[50] Sometimes the scorecard and PovcalNet agree, and some disageements stem from choices for which there is no clear right or wrong. And both the scorecard and PovcalNet make some mistakes. But on the whole, it is difficult to oppose better data, better price adjustments, correct time units, more cross-checks, and better documentation. Thus, the scorecard's $1.25/day estimates are generally to be preferred.

The rest of this appendix explains a mistake by the PAT in its derivation of the median line and documents—for the 15 countries in which both the scorecard and the PAT use the same survey data—the details of their implementation of the national, median, and $1.25/day poverty lines. For these 15 countries, the scorecard's implementation of the $1.25/day line is also compared with that of PovcalNet.

[50] PovcalNet also prioritizes comparability across countries and over time, and this sometimes leads to (for example) not adopting new-and-improved definitions.

A.1 Definition of the median line

The PAT derives the median line incorrectly and so over-estimates poverty rates. To avoid this, the 19 PATs using the median line would have to be reconstructed.

According to U.S. Congress (2004, 188 STAT 3930, section 14), a country's USAID "very poor" line is the higher of the median line or $1.25/day:

"The term 'very poor' means those individuals—

> (A) living in the bottom 50 percent below the poverty line established by the national government of the country in which those individuals live; or
>
> (B) living on less than the equivalent of $1[.25] per day."

The scorecard and PAT read the definition of "very poor" (part A) differently:[51]

Scorecard: *individuals* in the bottom 50 percent of [*individuals*] below the national line
PAT: *households* in the bottom 50 percent of [*households*] below the national line

While the law leaves the bracketed concept implicit, the first term is explicitly *individuals*, which can only mean *persons*, not *households*.[52] And given that the first term is *individuals*, the implicit term must also be *individuals*.

[51] For the case of Mexico, see Schreiner (2009b, p. 13) and IRIS Center (2010e, p. 3).
[52] IRIS Center (2005, p. 2) mis-paraphrases the law, replacing *individuals* with *households*: "According to that legislation, a *household* is classified as "very poor" if either: (1) the *household* is 'living on less than the equivalent of a dollar a day' [. . .] or (2) the *household* is among the poorest 50 percent of *households* below the country's own national poverty line." (italics added)

But the PAT derives the median line so as to divide households (not individuals) below the national line into two equal groups. For the example of Mexico, this means that the household-level poverty rate for the median line (20.1 percent) is half of the household-level poverty rate for the national line (40.2 percent, Tables A1 and A2).

In contrast, the scorecard derives the median line to divide people (not households) below the national line into two equal groups. In the example of Mexico, this means that the person-level rate for the median line (23.7 percent) is half the person-level rate for the national line (47.4 percent).

Both the scorecard and the PAT are constructed and applied using household-level definitions of poverty, and perhaps this is what led to confusion.[53] But the law is clear, and *applying* the median line at the household level does not mean that it must also be *derived* at the household level.

The PAT's implementation of the median line leads to over-estimates of poverty. Household-level rates for the median line derived in terms of households (20.1 percent in the example of Mexico, Table A1) always exceed household-level poverty rates for the median line derived in terms of people (19.2 percent for Mexico). This is because poorer households tend to have more members than less-poor households.

[53] Schreiner (2012, section 2.2) explains why the construction and application of the scorecard and the PAT use household-level definitions of poverty.

A.2 Documentation of country-level details

For each of the 15 countries in which the scorecard and PAT use data from the same national survey, this section documents poverty rates for the national, median, and $1.25/day lines. The differences among the scorecard, PAT, and PovcalNet are usually due to errors related with the implementation of poverty lines.

Cambodia

Cambodia's 2003/4/5 Socio-Economic Survey was in the field for 15 months (November 2003 to January 2005). The scorecard follows World Bank (2006, p. 18), Knowles (2006, p. 41), and Ministry of Planning (2006, p. 47) in using only the 11,993 households from the 12 months of 2004. The scorecard's person-level poverty rate for the national line of 34.7 percent matches these three sources (Schreiner, 2009a, Table A2).

The PAT uses 14,984 households from all 15 survey months. Its household-level poverty rate for the median line—Cambodia's "very poor" line—is 15.3 percent (IRIS Center, 2009a, Table A1), implying a household-level rate for the national line of 30.6 percent (Table A2). Using all 15 months is mistake not only because the official estimate uses only the 12 months of 2004 but also because consumption expenditure is seasonal, and three calendar months appear twice in the 15-month data. This pulls the overall poverty rate towards the poverty rates typical in these three months.

This rules out a head-to-head comparison for Cambodia, as it would require reconstructing the scorecard from scratch with the 15-month data.

The scorecard's person-level poverty rate for $1.25/day of 37.9 percent matches PovcalNet's 37.7 percent (Table A3).

Ecuador

The scorecard and PAT start with data on 13,581 households from the 2005/6 Living Standards Survey, fielded for 12 months starting November 2005. The PAT drops 149 households for whom expenditure is missing (IRIS Center, 2010a). Expenditure is never missing in the scorecard's data, but the scorecard drops 61 households marked as not having completed the survey (Schreiner, 2008a). Ecuador (*Instituto Nacional de Estadística y Censos*, 2007) drops 45 households without giving a reason. In sum, the PAT, the scorecard, and Ecuador use slightly different samples. Thus, even if they use the same poverty lines, their poverty rates will not match.

The USAID "very poor" line is the median line, for which the PAT incorrectly finds a household-level poverty rate of 14.8 percent. To match this, the scorecard:

- Uses Ecuador's single national line (the scorecard incorrectly used regional lines to adjust for differences in the cost of a food basket)
- Follows the PAT's incorrect derivation of the median line in terms of households
- Multiplies the single national line by 0.9728 to get a household-level poverty rate for the national line and for the median line that matches those of the PAT. This step is needed because of the two tools start with different samples

For $1.25/day, PovcalNet reports a person-level poverty rate of 9.1 percent. This differs from the scorecard's 2.4 percent, probably mostly because PovcalNet uses the December 2005 Labor Force Survey. The scorecard is to be preferred because of its use of data from an expenditure survey.

El Salvador

The scorecard, PAT, and PovcalNet use data from the 2008 Multi-Purpose Household Survey, fielded for the 12 months of 2008. The USAID "very poor" line is the median line, for which the PAT reports an incorrectly derived household-level poverty rate of 20.2 percent (IRIS Center, 2010b). This implies a household-level rate for the national line of 40.4 percent, which matches El Salvador's official household-level rate for the national line (40.0 percent, Dirección General de Estadística y Censos, 2009, p. 21),[54] and the scorecard's rate (39.9 percent, Schreiner, 2010a). To match the PAT exactly, the scorecard increases the national line in urban and rural regions by a factor of 1.00793.

For $1.25/day, the reported person-level poverty rates are 11.0 percent (PovcalNet) and 52.7 percent (scorecard). The discrepancy is partly due to documented differences in how PovcalNet and the scorecard adjust prices from 2005 to 2008. While the scorecard's figure seems too high, PovcalNet—unlike the scorecard—does not tell how it:

- Adjusts for urban/rural differences in prices
- Converts the PPP factor in units of *SVC* in El Salvador in 2005 per dollar in the USA to units of *dollars* in El Salvador in 2005 per dollar in the USA[55]

Most of the discrepancy is due to different conversions of the PPP factor from SVC to dollars. PovcalNet does not report what it did, but it may have taken the PPP factor as 1:1. If so, it would be incorrect, as a dollar buys more in El Salvador than in the USA. The scorecard documents how it derives its PPP factor (1.82:1), and while this may or may not be correct, it at least is transparent, so the scorecard's estimate is to be preferred.

[54] El Salvador does not report a person-level poverty rate for the national line. This is unique among countries, and it is probably serves to report a lower poverty rate.
[55] In 2008, El Salvador had been fully dollarized for four years, but the World Bank (2008) reports El Salvador's 2005 PPP factor in units of SVC.

Ethiopia

Ethiopia's data come from two surveys six months apart that covered the same 21,297 households. The Welfare Monitoring Survey collected poverty indicators from 4 July to 3 August 2004, and the Household Income, Consumption, and Expenditure Survey collected expenditure from 2 February to 5 March 2005.

Ethiopia does not have a national poverty line. Thus, the USAID "very poor" line is $1.25/day. The expenditure data can be considered to be in average prices in February of 2005, as food expenditure was collected for the month of February.[56] Both the scorecard (Schreiner and Chen, 2009) and the PAT (IRIS Center, 2009b) adjust for prices incorrectly, with the PAT acting as if expenditure is in average prices for all months in 2005, and the scorecard acting as if expenditure were in average prices from July 2004 to February 2005.

When the scorecard fixes its error, its person-level poverty rate for $1.25/day is 35.9 percent, lower than PovcalNet's 39.0 percent and the PAT's 39.8 percent. Given the similarity of these last two estimates, PovcalNet's undocumented deflators are probably the same incorrect ones as the PAT's. Thus, the scorecard's (corrected) estimate for $1.25/day is to be preferred.

[56] Food expenditure exceeds non-food expenditure in Ethiopia. It would be better to weight price units by the shares of food and non-food expenditure, but the data do not have this break-down.

Ghana

Data are from Ghana's 2005/6 Living Standards Survey, fielded for 12 months starting September 2005. Poverty lines and expenditure are in prices as of January 2006, adjusted across five regions. The scorecard's person-level poverty rate for the national line (28.5 percent, Schreiner and Woller, 2010a) matches the official one (Ghana Statistical Service, 2007, p. 9). At the household level, the scorecard's poverty rate for the national line is 18.9 percent.

The PAT's median line "identifies 14.2 percent of *households* [italics added] in the sample as 'very poor'" (IRIS Center, 2010c, p. 3). This implies a *household*-level rate for the national line of 28.4 percent, which is almost the same as the official *person*-level rate for the national line of 28.5 percent. At the same time, the PAT's \$1.25/day person-level poverty rate of 28.8 percent matches PovcalNet's 28.6 percent but not the scorecard's 25.3 percent.

In other words: the scorecard matches the official national line, as does the PAT (if the PAT avoided the mistake in its median line that it made in all other countries). But the PAT and PovcalNet—but not the scorecard—match for \$1.25/day. So either:

- The scorecard has the \$1.25/day line wrong, or
- The PAT and PovcalNet make the same errors in deflators and regional lines

The second explanation is more likely. The scorecard adjusts for price differences across poverty-line regions, but the PAT reports a single value for the \$1.25/day line, and PovcalNet is silent on this topic. Furthermore, the scorecard derives its \$1.25/day line in prices as of January 2006, but PovcalNet does not mention its units, and the PAT says its line is "in 2005 prices" (IRIS Center, 2010c, p. 2). This is likely a typo for "in 2006 prices". But p. 3 says that "the value of [\$1.25] at the time of the survey is GHC188,749 per capita per month", which implies the use of average prices while the Ghana survey was in the field, not prices as of January 2006.

The deflator that is needed to match the PAT's reported poverty rates is 1.1360, whereas the average deflators for the calendar year 2006 (and the average deflators while the Ghana survey was in the the field) are about 1.1090, and the correct deflator for January 2006 is 1.0503.

Thus, the \$1.25/day line derived by the PAT and PovcalNet have the wrong units or use different data than that used by the scorecard and Ghana's government. The PAT and PovcalNet also fail to adjust for price differences across regions. Thus, the scorecard's poverty rates (for all poverty lines) are preferred.

For the comparison here, the scorecard's \$1.25/day line is multiplied by $1.1360 \div 1.0503 = 1.0816$ in each region so that the resulting household-level poverty rate matches the PAT's 18.9 percent.

Guatemala

Schreiner and Woller (2010b) construct a scorecard with Guatemala's 2006 LSMS. To measure the accuracy of its estimates of changes in poverty rates out-of-sample/out-of-time, they also set up data from the 2000 LSMS (fielded July to September) and implement poverty lines for it.

The Guatemala PAT is constructed with the 2000 LSMS (IRIS Center, 2006). The PAT has no out-of-sample tests, so its accuracy cannot be compared head-to-head with the scorecard. Nevertheless, the implementation of poverty lines can be compared.

The scorecard's person-level poverty rate for the 2000 national line (56.1 percent) matches the official rate (56.2 percent, World Bank, 2003, p. 8). The PAT reports a household-level poverty rate for the median line of 23.0 percent. This implies a household-level poverty rate for the national line for the PAT of 46.0 percent, and this matches the scorecard's 45.8 percent. This suggests that the scorecard and the PAT use the same data and implement the national line correctly.

Still, the PAT's median line is incorrect for two reasons. First it is derived—as in all countries—in terms of households rather than people. Second, it is a single line for all of Guatemala, but there should be a distinct median line for each of Guatemala's 22 poverty-line regions. Although a single all-country line gives the same poverty rate as region-specific lines, it does not count the same households as "very poor", and so the two approaches are not equivalent.

For the $1.25/day line, the scorecard's person-level poverty rate is 5.4 percent, while PovcalNet's is 11.9 percent. Given PovcalNet's weak documentation, the reason for the difference is unknown, but it may reflect PovcalNet's:

- Incorrectly deflating prices as if the 2000 LSMS was fielded over the calendar year
- Not adjusting for price differences across poverty-line regions, or
- Using income rather than expenditure (CEDLAS and World Bank, 2012)

The scorecard's estimate of $1.25/day poverty is to be preferred because it is based on expenditure, it is known to use price units that correspond to when the national survey was in the field, and it is known to adjust for price differences across Guatemala's poverty-line regions.

Kenya

Data are from Kenya's 2005/6 Integrated Household Budget Survey, fielded for 12 months starting May 2005. The scorecard has $n = 12{,}644$ (Schreiner, 2011a), and the PAT has $n = 13{,}158$ (IRIS Center, 2010d). The sample sizes differ because the scorecard drops 514 households with missing values for all items in the questionnaire modules on demographics, assets, and housing and energy/water/sanitation. The PAT keeps these 514 households and imputes their missing values. Neither approach is wrong. Expenditure is available for all 13,158 households.

The scorecard has two errors, and both have been fixed for the comparisons here. First, it mistakenly dropped 20 households who report owning no assets. Second, it failed to adjust its \$1.25/day line for the fact that the person-weighted regional-price deflators do not average exactly 1.0.

Both the fixed scorecard and the PAT implement the national line corrrectly. This is known because—when looking at all surveyed households—both tools give household-level poverty rates for the national line of 38.3 percent, and the scorecard's person-level poverty rate for the national line (46.6 percent) matches the official rate (National Institute of Statistics, 2007, p. 44).

The fixed scorecard and the PAT also report the same household-level poverty rate for \$1.25/day. This implies that the scorecard's person-level rate for \$1.25/day of 46.1 percent is to be preferred to PovcalNet's 43.4 percent.

The PAT's imputations cannot be replicated, so, for the comparison here, the scorecard's \$1.25/day line is increased by a factor of 1.0195 in each region so that the poverty rate in its reduced sample matches the PAT's rate of 36.3 percent in the PAT's sample.

Malawi

Data are from the 2004 Integrated Household Survey, fielded for 12 months starting March 2004. Both the scorecard (Schreiner, 2011b) and PAT (IRIS Center, 2009c) implement the national line correctly. This is known because the scorecard's person-level poverty rate for the national line matches the official rate (52.4 percent, World Bank, 2007, p. 4). Furthermore, the scorecard's household-level rate for the national line (43.6 percent) matches the 43.8 percent that is implied by the PAT's reported household-level rate for its incorrectly derived median line of 21.9 percent.

Both the scorecard and PAT have errors in the $1.25/day USAID "very poor" line. The PAT does not adjust for regional-price differences, and the scorecard does not account for the fact that the regional-price deflators do not average to exactly 1.0. The scorecard also incorrectly derives the $1.25/day line in average prices while the survey was in the field rather than prices as of February/March 2004 (the units in which expenditure is supplied by Malawi's National Statistical Office).

For the comparison here, correcting the scorecard's two errors and replicating the PAT's error gives a household-level poverty rate for the $1.25/day line of 64.7 percent, matching the PAT's 64.8 percent.

The corrected scorecard matches PovcalNet's person-level poverty rate for the $1.25/day line (74.2 percent and 73.9 percent).

Mexico

Data for 29,468 households in the August 2008 National Household Income and Expenditure Survey are provided by the *Instituto Nacional de Estadística, Geografía e Informática*. The PAT drops 282 households—without saying why—and has $n = 29{,}186$ (IRIS Center, 2010e). This sample difference probably explains the differences in scorecard/PAT poverty rates below.

The scorecard and the PAT document the use of the same national poverty line. Furthermore, it is known that the two tools implement this line correctly because the scorecard's person-level poverty rate matches the official one (47.4 percent, CONEVAL, 2009, p. 13). The household-level poverty rate of 20.1 percent for the PAT's incorrectly derived median line implies a household-level rate for the national line of 40.2 percent, while the scorecard's is 40.7 percent. The difference is probably due to the 282 households dropped by the PAT.

For the comparisons using the USAID "very poor" line here, the scorecard's median line is made to replicate the PAT's mistake and then increased by a factor of 0.9946 to account for the 282 dropped households.

For \$1.25/day, the scorecard's person-level poverty rate of 2.2 percent differs from PovcalNet's 1.2 percent. PovcalNet may define expenditure differently than Mexico does, and its adjustment for urban/rural price differences is known to differ from the official one (CEDLAS and World Bank, 2012). Thus, the scorecard's estimate is to be preferred.

Nepal

Data are from the 2003/4 Nepal Living Standards Survey, fielded for 12 months starting April 2003. The PAT does not test accuracy out-of-sample, so accuracy is not compared head-to-head with the scorecard. Still, poverty lines and rates are compared.

For the USAID "very poor" line of $1.25/day, the PAT's household-level poverty rate is 52.6 percent (IRIS Center, 2009d), and the scorecard's is 47.8 percent (Caire *et al.*, 2009).[57] Which is correct? The scorecard's person-level rate for the national line is 30.8 percent, matching the official 30.9 percent (Central Bureau of Statistics, 2005, p. 2). Thus, the scorecard implements the national line correctly. The scorecard's household-level rate for the national line of 25.9 percent differs from the 29.8 percent that is implied by the PAT's reported household-level rate for the median line of 14.9 percent.

In addition, PovCalNet reports a *person-level* $1.25/day poverty rate of 53.1 percent, which—given that the PAT reports a *household-level* $1.25/day poverty rate of 52.6 percent—must be closer to the scorecard's *person-level* rate of 53.6 percent than to the PAT's unreported person-level rate.

The differences stem from two mistakes by the PAT. First, the PAT does not adjust the $1.25/day line for price differences across Nepal's poverty-line regions. Second, the PAT uses the wrong deflator to adjust the 2005 PPP factor to units of average prices while the 2003/4 NLSS was in the field. This is known because both tools report using a 2005 PPP factor of 26.467, but the PAT uses a deflator of 0.95191 (without reporting the source) while the scorecard uses a deflator of 0.9160 (reporting the source of its $1.25/day line—and implicitly its deflator—as PovcalNet's Prem Sangraula). Given monthly Consumer Price Indexes for Nepal (base January/February 2002 = 100), the average CPI while the survey was in the field was 110.48,[58] and the average CPI in 2005 was 120.2863. The scorecard's resulting deflator of 0.9185 is close to PovcalNet's 0.9160 but far from the PAT's 0.95191.

For $1.25/day, the scorecard and PovCalNet use the same line, yet their person-level poverty rates do not match (53.6 percent versus 53.1). Given that the scorecard is known to implement the national line correctly, the most likely explanation is PovcalNet's use of grouped data instead of household-level data (communication with Prem Sangraula). Thus, the scorecard's estimate of $1.25/day poverty is to be preferred.

---

[57] Schreiner (2013a) updates the Nepal scorecard by Caire *et al.* (2009) with data from the 2010 NLSS. The older scorecard is used here for comparability with Nepal's PAT.

[58] This CPI splices together several series from issues of Nepal's *Quarterly Economic Bulletin*, `red.nrb.org.np/publications/economic_bulletin`, retrieved 10 April 2013. The resulting CPI series for Nepal—as well as the monthly CPIs used for other scorecards—are available on request.

Nicaragua

The data provided by Nicaragua's *Instituto Nacional de Información de Desarrollo* for the 2005 Living Standard Measurement Survey (fielded from July to October) covers 6,882 households. The scorecard drops 30 households who did not complete both survey rounds (Schreiner and Woller, 2010c)[59], while the PAT drops 35 households without reporting the reason.

Both the scorecard and PAT report a household-level poverty rate for the national line of 39.2 percent, and the scorecard's person-level rate for this line (48.4 percent) matches the official 48.3 percent (*Instituto Nacional de Información de Desarrollo*, 2007, p. ix). Thus, both tools implement the national line correctly.

The USAID "very poor" line is the median line. While the scorecard's person-level poverty rate for the median line of 23.8 percent is correctly half the person-level rate for the national line of 48.4 percent,[60] the PAT's household-level poverty rate for the median line of 19.6 percent is incorrect—as in all countries—because it derived in terms of households rather than people (IRIS Center, 2011a). For the comparison here, the scorecard matches the PAT's household-level rate for the median line by replicating its error.

The reported person-level poverty rates for $1.25/day are the same for the scorecard (11.8 percent, Schreiner and Woller, 2010c, p. 70) and PovcalNet (11.9 percent). The scorecard, however, is in error in that it does not adjust for the fact that the regional price deflators do not average to 1.0. Given the similarity of its reported rate, PovcalNet probably makes the same mistake. When the scorecard is corrected, its person-level poverty rate is 11.3 percent, and this estimate is to be preferred.

---

[59] Schreiner (2013b) updates Schreiner and Woller's (2010c) Nicaragua scorecard with 2009 data. The older scorecard is used here for comparability with Nicaragua's PAT.

[60] In countries with many poverty-line regions (Nicaragua has 34, urban and rural for each of 17 departments), the scorecard's person-level poverty rate for the median line may be a little less than half the person-level rate for the national line (in Nicaragua, 23.8 is less than half of 48.4 percent). This is not an error, however, because it is due to a conflict between the definition of the median line (which is in terms of people and thus splits the members of one household into poor and non-poor in each poverty-line region) and the definition of poverty status (which says that all members of a household have the same poverty status). In Nicaragua, the scorecard follows the definition of household status and counts all the members in the household in which the median person lives as non-poor. Accumulated over many poverty-line regions, makes the person-level poverty rate too low.

Nigeria

Data are from the 2003/4 Nigeria Living Standards Survey, fielded for 12 months starting September 2003. Both the scorecard and the PAT implement the national line correctly. The scorecard's person-level poverty rate of 54.5 percent for the this line (Schreiner, 2008b) matches the official 54.4 percent (National Bureau of Statistics, 2005, p. 21). Also, the scorecard's household-level rate of 44.8 percent for this line is close to the 45.2 percent rate implied by the PAT's household-level 22.6-percent rate for its incorrectly derived median line (IRIS Center, 2011b).

The USAID "very poor" line is $1.25/day. Because the scorecard was constructed before the advent of 2005 PPP factors, it was not calibrated to $1.25/day 2005 PPP but rather to $1.08/day 1993 PPP. For the comparisons here, the scorecard adopts the PAT's $1.25/day line of NGN81.46. This yields a household-level poverty rate in the data of 58.0 percent, matching the PAT's reported 58.2 percent.

But the PAT's $1.25/day line is in error. Expenditure in the data is in prices as of January 2004, but the PAT mistakenly converts the $1.25/day line to prices in units as of "the time of the survey" (IRIS Center, 2011b, p. 3). Given the 2005 PPP factor of 78.583, and average CPIs (source undocumented) for September 2003 to August 2004 of 82.9291 and for all of 2005 of 100, the PAT's $1.25/day line is 1.25 x 78.583 x 82.9291 ÷ 100 = NGN81.46.

The January 2004 CPI (base November 2009 = 100) is 55.4610, and the average CPI for 2005 is 66.5751.[61] Thus, the $1.25/day line should be 1.25 x 78.583 x 55.4610 ÷ 66.5761 = NGN81.83. This gives a household-level poverty rate of 58.3 percent and a person-level rate of 68.5 percent.

PovcalNet reports a person-level rate for $1.25/day of 63.1 percent but does not document its deflators. In the data, a person-level rate of 63.1 percent implies a deflator of 0.7460, which is close to 0.7344, the ratio of the average CPI for 2003 (48.8950) to the average CPI for 2005 (66.5751). PovcalNet sometimes ignores the actual dates of a country's national expenditure survey, assuming instead that it was fielded in the 12 months of a calendar year. This seems to be the case in Nigeria. In sum, PovcalNet probably uses the wrong time units for its $1.25/day line, so the scorecard's estimate here is to be preferred.

[61] Nigeria's monthly CPIs are derived from `cenbank.org/Functions/export.asp?tablename=InflationRates`, retrieved 20 October 2013.

For Nigeria, the scorecard and the PAT differ in an aspect unrelated to poverty lines and rates. In the 2003/4 NLSS data, 4,045 households do not have a record in the data file pertaining to the ownership of durable consumer assets. The PAT assumes that asset-ownership is missing for these households and so excludes all types of asset ownership from its 19 poverty indicators (IRIS Center, 2011b). The scorecard assumes that the absence of a record means that the household did not own any of the assets (Schreiner, 2008b).[62] A few clues are consistent with the scorecard's assumption:

- Nigeria as a whole is very poor, and most poor households are rural farmers for whom not owning any of the covered assets is plausible, as many of the assets are large (such as refrigerators or motor vehicles) or run on electricity from the grid
- Households with no asset records have lower average daily per-adult-equivalent expenditure than others (NGN93 versus 134). For the national line, the household-level poverty rate for no-asset households is 66 percent, versus 39 percent for others
- Most asset indicators in the scorecard (television, stove, mattress/bed, and radio) are small, and only one requires a connection to the electrical grid. Moreover, the asset indicators' points make sense and are not trivially small. When the scorecard is re-estimated without no-asset households, the points for the asset indicators shrink a lot, inconsistent with the idea that owning assets signals less poverty

This suggests that the households with missing records probably do not have any of the covered assets, rather than that their asset data is missing. As a testament to the strength of the "flat maximum", however, this does not translate into any unusual differences in relative accuracy between the two tools in out-of-sample/in-time tests.

[62] A few other countries omit no-asset households from the asset-ownership data file, although the share of such households is usually smaller than Nigeria's 20 percent.

Peru

IRIS Center (2011c) constructs a PAT for Peru with data from the 2009 National Household Survey (ENAHO), fielded from January to December. Schreiner (2012) constructs a scorecard not with the 2009 ENAHO but rather with the 2010 ENAHO. To measure accuracy both out-of-sample/in-time and out-of-sample/out-of-time, the scorecard based on the 2010 construction sample is tested against both the 2010 validation sample and the entire 2009 ENAHO.

The scorecard's person-level poverty rate for the national line with the 2009 ENAHO is 34.8 percent, matching *Instituto Nacional de Estadística e Informática* (2011, p. 35). Thus, the scorecard is known to implement the national line correctly.

At the household-level, the scorecard's rate for the national line (29.0 percent) differs from the PAT's 31.8 percent. The PAT is in error; it uses a single value rather than distinct lines in Peru's 50 poverty-line regions (urban and rural in 25 departments).

The USAID "very poor" line is the median line. The PAT's median line is in error in that it is derived incorrectly (based on households) and in that it uses one all-Peru line instead of 50 region-specific lines. The scorecard did not derive median lines for the 2009 ENAHO, so for the comparisons here, it replicates the PAT's errors.

For \$1.25/day, PovcalNet's person-level rate is 5.5 percent, and the scorecard's is 0.6 percent. Both use the same documented deflator to take the 2005 PPP factor from average 2005 prices to average 2009 prices, so the difference must stem from PovcalNet's use of income rather than expenditure and from PovcalNet's probable use of a single line for all of Peru rather than 50 region-specific lines. The scorecard's estimate is preferred because it is better to measure poverty with expenditure and because it is better to follow Peru's use of region-specific lines.

Rwanda

Data are from the 2005/6 Integrated Household Living Standards Survey, fielded for 12 months starting October 2005. Rwanda's *Institut National de la Statistique* provides expenditure data in prices as of January 2001, along with deflators that bring expenditure to January 2006 for five poverty-line regions.

The scorecard's person-level poverty rate of 56.9 percent (Schreiner, 2010b) matches the official one (*Institut National de la Statistique*, 2006, p. 3), establishing that the scorecard implements the national line correctly.

The scorecard's household-level poverty rate for the national line is 54.1 percent, matching the PAT's reported 53.9 percent. The PAT's reported household-level rate for its (incorrectly derived) median line is 26.9 percent (IRIS Center, 2011d).

The USAID "very poor" line is \$1.25/day. The scorecard defines this line in prices as of January 2006 (the unit of the measure of expenditure) and adjusts it across poverty-line regions to get a person-level rate of 71.7 percent and a household-level rate of 69.5 percent. The PAT mistakenly defines the \$1.25/day line in units of average prices for 2006, giving a household-level rate of 72.9 percent.

For the comparisons here, the scorecard's \$1.25/day line is redefined to follow the PAT's error. The resulting household-level rate of 72.1 percent, however, still does not match the PAT's 72.9 percent. This reason for this discrepancy is unknown, as both tools use the \$1.25/day line documented by the PAT, the same regional deflators, and the same measure of expenditure. The discrepancy is resolved by increasing the scorecard's \$1.25/day line in each region by a factor of 1.018.

PovcalNet's person-level poverty rate for \$1.25/day is 72.1 percent, close to the scorecard's 71.7 percent. The reason for the small difference is unknown.

Tanzania

Data for the 2007 Household Budget Survey from the National Bureau of Statistics came with two household-roster files and no documentation. One file included all household members, but the other file omitted some members of some households.[63] Both the scorecard (Schreiner, 2011c) and the PAT (IRIS Center, 2011e) used the wrong data file. When the error was discovered in 2011, the scorecard was fixed, and the PAT developer was alerted. The Tanzania PAT has since been decommissioned (USAID, 2013) and is no longer available on `povertytools.org`.

The error prevents a scorecard/PAT comparison. The PAT's household-level poverty rate for Tanzania's "very poor" line ($1.25/day)[64] is 43.0 percent, versus 61.3 percent for the scorecard using the correct data (after also correcting an error in the scorecard's application of regional deflators to the $1.25/day line).

For the national line, the scorecard matches the official person-level poverty rate of 33.6 percent (National Bureau of Statistics, 2009, p. 49). For $1.25/day at the person level, the scorecard reports 71.3 percent, versus 67.9 percent for PovcalNet. The reason for this discrepancy is unknown, but the scorecard is probably more accurate, given that it is known to have implemented the national line correctly.

[63] PAT developers were aware of a discrepancy (IRIS Center, 2011e, p. 3), noting that using a summary measure of household size found elsewhere in the data from the National Bureau of Statistics produced a person-level poverty rate for the national line 33.9 percent, close to the official 33.6 percent.

[64] The PAT does not report a median line nor poverty rate for Tanzania.

Figure 1: PAT for Peru

Client Assessment Survey - Peru

Interviewer: **Text in bold** should be read aloud. *Text in italics* are instructions and should not be read aloud.

You should use probing questions if necessary to elicit responses to all questions. If, however, a response is still not forthcoming, the following codes should be used: 99 - not applicable; 98 - no response given.

Fill out the information below before the survey begins. Do not ask the respondent for this information.

Date of Interview (dd-mm-yyyy)	
Interviewer (code)	
Branch (code)	
Region	*1 □ Costa Norte* *2 □ Costa Centro* *3 □ Costa Sur* *4 □ Sierra Norte* *5 □ Sierra Centro* *6 □ Sierra Sur* *7 □ Selva* *8 □ Lima Metropolitana*
Client Location	*0 □ Urban* *1 □ Rural*
Months in Program	
Client or ID #	

Quality Control Checks
Field Supervisor Date_____Initials
Headquarters Date_____Initials
Data Processor Date_____Initials

Hello. My name is ___. I work for the organization ____. We are conducting a survey to learn a little bit more about the clients we work with. My records indicate that [name**] is the main point of contact between [**organization**] and your household. May I please speak to [**name**]?**

If person is desired respondent, read only the instructions marked 2.

If person is NOT desired respondent, read both 1 and 2 when desired respondent is located.

1. **Hello. My name is ___. I work for the organization ____. We are conducting a survey to learn a little bit more about the clients we work with. My records indicate that you are the point of contact between [**organization**] and your household.**

2. **The interview should only take about 20 minutes and your answers will be put together with answers from other households. All of your answers are completely confidential and your name will not be given with your answers. Are you willing to answer these questions today?**

After he/she agrees, proceed with the text below.

First, I would like to ask you about your household. Let me tell you what we mean by 'household.' For our purposes today, members of a household are those that usually live and eat together in this dwelling. It should include anyone who has lived in your house for 6 or more of the last 12 months, as well as the person you identify as the head of household if he or she has been absent for more than 6 of the last 12 months and infants under 6 months of age who normally live and eat here. Do you have any questions about that?

Answer any questions the respondent has before proceeding.

Now I would like you to identify each person in your household and answer some basic questions about each person. Let's start with the names of each person in your household. Shall I identify you as [name**]?**

If the respondent is reluctant to provide his or her name or those of others in the household, record relationships instead (ex: Respondent, Husband, etc).

Use row 1 for respondent.

Are you the head of the household or is someone else?

If not the respondent, record the Head of Household's name next, then continue filling in column A with each household member before asking questions in the remaining columns.

1.	A. **Household Member's Name**	B. **Is [**NAME**] female or male?**	C. **What is the relationship of [**NAME**] to [**HOUSEHOLD HEAD**]?**	D. **How old is [**NAME**]?**	E. *Ask only if age 12 or older* **What is [**NAME**]'s marital status?**	F. *Ask only if age 3 or older* **What is the last level of education that [**NAME**] has obtained?**
		Female .... 0 *Male 1*	*Head ..... 1* *Spouse/Partner ..... 2* *Son/Daughter ..... 3* *Son-in-law/Daughter-in-law .. 4* *Grandchild ..... 5* *Parents/Parents-in-law..... 6* *Other Relative ..... 7* *Domestic Worker..... 8* *Pensioner ..... 9* *Other, Non-relative 10*	*(complete years)*	*Under age 12.....99* *Lives with a partner.....1* *Married.....2* *Widowed.....3* *Divorced.....4* *Separated.....5* *Single.....6*	*Under age 3.....99* *None .....1* *Preschool.....2* *Primary School, Incomplete .....3* *Primary School, Completed.....4* *Secondary School, Incomplete.....5* *Secondary School, Completed..... 6* *Technical School, Incomplete..... 7* *Technical School, Completed..... 8* *University, Incomplete .....9* *University, Completed .....10* *Graduate School.....11*
1)						
2)						
3)						
4)						
5)						
6)						
7)						
8)						
9)						
10)						
11)						
12)						
13)						
14)						
15)						

Now, I would like to ask you a few questions about your home.

2. **How many bedrooms does this dwelling have?** ☐ *enter number*

3. **What is the predominant material used in the roof of your dwelling?**
   - *1* □ *Reinforced Concrete*
   - *2* □ *Wood*
   - *3* □ *Tile*
   - *4* □ *Tin sheet, fiber cement, or similar material*
   - *5* □ *Cane or reed mat with mud*
   - *6* □ *Reed mat*
   - *7* □ *Straw, palm leaves, etc.*
   - *8* □ *Other*

4. **What fuel is usually used in your household to cook food?**
   - *1* □ *Electricity*
   - *2* □ *Liquefied Petroleum Gas*
   - *3* □ *Natural Gas*
   - *4* □ *Kerosene*
   - *5* □ *Coal*
   - *6* □ *Firewood*
   - *7* □ *Other*
   - *8* □ *Do Not Cook*

Next, I would like to ask you about a few items that members of your household may own and use.

5. **Does your household own a gas stove?**
   - *0* □ *No*
   - *1* □ *Yes*

6a. **Does your household own a radio?**
   - *0* □ *No*
   - *1* □ *Yes*

6b. **How many radios does your household own?** ☐ *number or "0" if 6a response was "0"*

7a. **Does your household own a color TV?**
   - *0* □ *No*
   - *1* □ *Yes*

7b. **How many color TVs does your household own?** ☐ *number or "0" if 7a response was "0"*

8a. **Does your household own a refrigerator or freezer?**
   - *0* □ *No*
   - *1* □ *Yes*

8b. **How many refrigerators or freezers does your household own?** ☐ *number or "0" if 8a response was "0"*

9a. **Does your household own a car, van or pick-up truck?**
   - *0* □ *No*
   - *1* □ *Yes*

9b. **How many cars, vans or pick-up trucks does your household own?** ☐ *number or "0" if 9a response was "0"*

Look over the survey to see if you have missed any questions, then end the interview.

Those are all the questions I need to ask you today. Thank you for your time and effort in completing this survey.

Figure 2: Poverty scorecard and backpage worksheet for Peru

Entity	**Name**	**ID**	**Date** (DD/MM/YY)	
Participant:			Joined:	
Field agent:			Today:	
Service point:			Household size:	

Indicator	Response	Points	Score
1. How many members does the household have?	A. Seven or more	0	
	B. Six	7	
	C. Five	12	
	D. Four	17	
	E. Three	22	
	F. Two	27	
	G. One	34	
2. In the past week, how many household members ages 14 or older did any work? (not counting household chores)	A. One or none	0	
	B. Two	2	
	C. Three	6	
	D. Four or more	9	
3. What is the highest educational level that the female head/spouse completed?	A. None, pre-school, or kindergarten	0	
	B. Grade school (incomplete)	3	
	C. Grade school (complete), or high school (incomplete)	4	
	D. No female head/spouse	6	
	E. High school (complete), or non-university superior (incomplete)	7	
	F. Non-university superior (complete), or higher	13	
4. How many rooms are used only as bedrooms?	A. None	0	
	B. One	2	
	C. Two	4	
	D. Three or more	8	
5. What is the main material of the exterior walls?	A. Mud, matting, wattle and daub, adobe, stone with mud, or other	0	
	B. Wood, stone, stone blocks with mortar or cement, or brick or cement blocks	4	
6. What fuel does the household most frequently use for cooking?	A. Charcoal, kerosene, or other	0	
	B. Firewood	3	
	C. Gas (LPG or natural), electricity, or does not cook	7	
7. Does the household have a refrigerator/freezer?	A. No	0	
	B. Yes	3	
8. Does the household have a blender?	A. No	0	
	B. Yes	6	
9. How many color televisions does the household have?	A. None	0	
	B. One	5	
	C. Two or more	9	
10. Does the household have a cellular telephone?	A. No	0	
	B. Yes	7	
		Score:	

Back-page Worksheet:
Household Roster and Work Status

At the start of the interview, read to the respondent: *Please tell me the names and ages of all* household members*, that is all people—regardless of blood relationship—who stay or live permanently in the same residence, who share their main meals, and who cooperate together to fulfill their other basic needs. This includes whomever the household head thinks it should include. A household may have just one person. Do not forget absent members or newborns. Do not count live-in domestic servants nor lodgers.*

Write the names and ages all household members. For each member 14-years-old or older, ask whether he/she did any work in the past week (not including household chores).

See the "Guidelines to the Interpretation of Indicators" for more detail about the definitions of *work*, *past week*, and *household member*.

Count the number of household members, write it next to "Household size:" in the scorecard header, and mark the corresponding response to Question 1. Count the number of household members who work, and mark the response for Question 2.

Name of household member	Age	If <name> is 14-years-old or older, then ask: In the past week, did <name> do any work? (not counting household chores)
1.		No Yes
2.		No Yes
3.		No Yes
4.		No Yes
5.		No Yes
6.		No Yes
7.		No Yes
8.		No Yes
9.		No Yes
10.		No Yes
11.		No Yes
12.		No Yes
13.		No Yes
14.		No Yes
15.		No Yes
Total members:		**Total workers:**

Figure 3: Peru PAT in poverty-scorecard format

Indicator	Response	Points	Score
1. How many members does the household have?	Let the number of household members be h. Then points are h x (h x 0.0104 – 0.2170) + 8.7703		
2. What is the age of the household head?	Let the age be a. Then points are a x (0.0002 – a x 0.0000)		
3. Does the household live in a rural area?	A. Yes	–0.1751	
	B. No	0	
4. In what region does the household live?	A. Sierra Sur	–0.2305	
	B. Costa Sur	–0.1885	
	C. Sierra Norte	–0.1805	
	D. Sierra Centro	–0.1438	
	E. Costa Norte	–0.1118	
	F. Costa Centro	–0.0628	
	G. Selva	0	
	H. Lima Metropolitana	0.0522	
5. What is the dependency ratio?	Divide the number of household members 14 or younger or 65 or older by the number who are 15 to 64, and multiply the result by –0.0811		
6. Does the household head live with a partner?	A. Yes	–0.0712	
	B. No	0	
7. What is the highest level of education completed by the household head?	A. None	–0.1761	
	B. Other	0	
	C. Technical school	0.1204	
	D. University	0.2950	
8. What share of household members have no education?	Divide the number of members with no education by the total number of members, and multiply that by –0.0811		
9. How many rooms are in the dwelling?	Multiply the number of rooms by 0.0495		
10. What is the main construction material of the roof of the dwelling?	A. Straw, palm leaves, or etc.	–0.1889	
	B. Other	0	
	C. Reinforced concrete	0.1540	
11. Is firewood the main cooking fuel?	A. Yes	–0.0927	
	B. No	0	
12. Does the household own any gas stoves?	A. No	0	
	B. Yes	0.1200	
13. How many radios does the household own?	Multiply the number of radios by 0.0430		
14. How many color televisions does the household own?	Multiply the number of TVs by 0.1535		
15. How many refrigerators or freezers does the household own?	Multiply the number of refrigerators and freezers by 0.1666		
16. How many cars, vans, or pick-up trucks does the household own?	Multiply the number of cars, vans, or pick-up trucks by 0.1801		
		Score:	

Note: The PAT's reported coefficient of 0.0000 for indicator 2 must be missing some decimal places. The PAT does not say how many points to give indicator 5 if the household has no members ages 14 to 64. The presentation here combines the points for the constant/intercept term with those of indicator 1. More points means more expenditure.

Figure 4: Inclusion, undercoverage, leakage, and exclusion when using a poverty-measurement tool for targeting

True poverty status	Classification: Targeted	Classification: Not targeted
Poor	**Inclusion** Poor and Correctly Targeted	**Undercoverage** Poor but Mistakenly Non-targeted
Non-poor	**Leakage** Non-poor but Mistakenly Targeted	**Exclusion** Non-poor and Correctly Non-targeted

Table 1: For each country with a scorecard or PAT, the year of the national-survey data from which the tool is constructed, whether both tools exist and are directly comparable because they use the same data, the number of poverty lines supported, whether accuracy is tested out-of-sample, and out-of-sample/in-time bias and precision for estimates of poverty rates

	Year of data		Directly	Poverty lines		Out-of-sample?		Bias		Precision (α)	
Country	Card	PAT	comparable?	Card	PAT	Card	PAT	Card	PAT	Card	PAT
Afghanistan	2007/8		No PAT	6		X		1.3		0.95	
Albania		2002	No scorecard		2						
Azerbaijan		2002	No scorecard		2						
Bangladesh	2010	2004	Different data	9	5	X		-0.3		0.86	
Benin	2010		No PAT	6		X		0.4		0.94	
Bolivia	2007	2005	Different data	8	2	X	X	-1.2	-1.7	1.04	1.04
Bosnia and Herzegovina		2004	No scorecard		2		X		-1.2		1.37
Brazil	2008		No PAT	9		X		0.2		0.87	
Burkina Faso	2003		No PAT	7		X		-0.1		0.97	
Cambodia	2004	2003/4/5	Different data	9	2	X	X	0.3	-3.0	0.89	0.86
Cameroon	2007		No PAT	7		X		-2.3		1.15	
Colombia	2009	2003	Different data	18	1	X		0.3		1.16	
Côte d'Ivoire	2008		No PAT	8		X		1.1		0.89	
Dominican Republic	2007		No PAT	10		X		-1.6		0.80	
Ecuador	2005/6	2005/6	Yes	7	5	X	X	-0.3	0.2	0.77	1.06
Egypt	2004/5		No PAT	7		X		-0.1		0.90	
El Salvador	2008	2008	Yes	7	3	X	X	0.2	-0.5	0.88	1.26
Ethiopia	2004/5	2004/5	Yes	4	2	X		2.4		0.98	
Ghana	2005/6	2005/6	Yes	8	6	X	X	1.5	-0.8	0.82	1.03
Guatemala	2006	2000	See table note	8	2	X		1.5		0.87	
Haiti	2001		No PAT	1							
Honduras	2007		No PAT	6		X		1.1		0.93	
India	2009/10	1998	Different data	19	1	X		0.5		1.01	

Table 1 (continued)

Country	Year of data		Directly	Poverty lines		Out-of-sample?		Bias		Precision (α)	
	Card	PAT	comparable?	Card	PAT	Card	PAT	Card	PAT	Card	PAT
Indonesia	2010	2002	Different data	9	5	X		-0.5		1.01	
Jamaica		2000	No scorecard		2						
Jordan	2006		No PAT	9		X		0.2		1.29	
Kazakhstan		2004	No scorecard	2							
Kenya	2005/6	2005/6	Yes	6	3	X	X	1.0	0.0	0.85	1.23
Kosovo		2000	No scorecard		2		X		-1.9		1.15
Liberia		2008	No scorecard		5						
Madagascar		2001	No scorecard		2		X		-2.3		1.08
Malawi	2004/5	2004/5	Yes	7	2	X	X	0.0	0.3	0.85	0.94
Mali	2001		No PAT	5		X		-5.2		0.98	
Mexico	2008	2008	Yes	8	5	X	X	-0.8	0.4	1.03	1.61
Morocco	2007		No PAT	8		X		0.0		0.93	
Mozambique	2008/9		No PAT	6		X		-2.1		1.50	
Myanmar	2009/10		No PAT	7		X		1.5		0.90	
Namibia	2009/10		No PAT	8		X		-0.7		0.90	
Nepal	2010	2003/4	See table note	11	2	X		1.5		0.89	
Nicaragua	2009	2005	Yes	8	5	X	X	0.2	-0.6	0.77	1.13
Niger	2007/8		No PAT	8		X		3.4		1.14	
Nigeria	2003/4	2003/4	Yes	8	5	X	X	-0.8	0.7	1.26	1.13
Pakistan	2005/6		No PAT	9		X		0.6		0.92	
Palestine	2007	2007	Different data	7	2	X		-0.1		0.90	
Paraguay	2011	2000/1	Different data	7	2	X	X	0.3	1.4	0.99	2.36
Peru	2010	2009	Yes	15	5	X	X	-4.7	3.3	1.05	0.77
Philippines	2004		No PAT	8		X		0.7		0.57	
Romania	2007		No PAT	8		X		-0.8		1.13	
Russia	2007		No PAT	3		X		0.9		0.27	
Rwanda	2005/6	2005/6	Yes	7	5	X	X	-1.0	0.0	0.77	1.00

Table 1 (continued)

Country	Year of data		Directly comparable?	Poverty lines		Out-of-sample?		Bias		Precision (α)	
	Card	PAT		Card	PAT	Card	PAT	Card	PAT	Card	PAT
Senegal	2005/6	2009	Different data	10	5	X		2.0		1.16	
Serbia		2007	No scorecard		1		X		-0.2		1.12
Sierra Leone	2003/4		No PAT	7		X		1.2		0.73	
South Africa	2005/6		No PAT	7		X		-3.6		1.77	
Sri Lanka	2006/7		No PAT	10		X		0.0		0.96	
Syria	2006/7		No PAT	7		X		0.3		0.95	
Tajikistan		1999	No scorecard		1						
Tanzania	2007	2007	Different data	7	5	X					
Timor-Leste	2007	2001	Different data	7	1	X		1.0		1.08	
Uganda	2009/10	2004	Different data	8	5	X		0.3		0.97	
Vietnam	2006	1997/8	Different data	7	2	X		0.6		0.57	
Yemen	2005/6		No PAT	7		X		1.3		1.10	
Zambia	2010		No PAT	16		X		0.0		0.79	

A blank means "does not exist".

Bias is out-of-sample/in-time for the USAID "very poor" poverty line in terms of percentage points.

The scorecards for Haiti and the Philippines are also branded as PATs.

For Palestine, the scorecard covers both the West Bank and the Gaza Strip, and the PAT covers only the West Bank.

In Bangladesh, Liberia, Kazakhstan, Senegal, and Uganda, the PAT is based on data from a special-purpose survey by USAID.

The PAT for India covers only rural areas of the states of Bihar and Uttar Pradesh.

For Guatemala and Nepal, an older scorecard uses the same data as the PAT, but the two tools are not directly comparable because the PAT does not report out-of-sample/in-time accuracy.

Table 2: Look-up table to convert scores to poverty likelihoods, Peru scorecard

	Poverty likelihood (%)							
	National					Intl. 2005 PPP		
Score	Food	100%	150%	200%	Median	$1.25	$2.50	$3.75
0–4	73.7	100.0	100.0	100.0	83.5	45.4	72.6	100.0
5–9	70.6	98.5	99.5	100.0	78.8	12.3	66.4	93.7
10–14	57.5	95.8	99.4	100.0	72.2	4.7	47.4	90.1
15–19	43.3	91.7	99.4	100.0	58.2	2.2	40.3	80.5
20–24	39.7	84.5	96.7	99.6	53.5	2.1	35.2	72.6
25–29	27.5	77.0	94.8	99.3	46.1	1.9	25.1	61.5
30–34	17.8	66.9	90.7	98.1	32.3	1.0	16.7	48.8
35–39	9.5	52.0	85.3	95.4	22.4	0.4	8.9	34.4
40–44	4.8	38.9	76.8	93.6	18.4	0.3	4.8	23.6
45–49	1.4	26.5	63.9	83.9	8.0	0.1	1.9	11.8
50–54	0.6	16.8	53.6	77.2	4.3	0.0	0.7	5.2
55–59	0.0	8.1	38.5	67.9	2.3	0.0	0.0	2.3
60–64	0.0	3.6	25.8	53.3	1.0	0.0	0.0	1.2
65–69	0.0	1.5	14.5	38.3	0.3	0.0	0.0	0.3
70–74	0.0	0.7	6.5	20.2	0.2	0.0	0.0	0.0
75–79	0.0	0.0	2.1	8.3	0.0	0.0	0.0	0.0
80–84	0.0	0.0	0.0	4.5	0.0	0.0	0.0	0.0
85–89	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
90–94	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
95–100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 3: For each country in which the scorecard and PAT can be directly compared, the USAID "very poor" line, out-of-sample/in-time bias and precision for estimates of a group's poverty rate, and targeting accuracy in terms of the scorecard-minus-PAT differences in inclusion and exclusion

	USAID	Accuracy of estimated poverty rates					Difference in targeting accuracy, scorecard minus PAT				
	"very poor"	Bias			Precision (α)						
Country	line	Card	PAT		Card	PAT	Inclusion		Exclusion		Hit rate
Ecuador	Median	-0.3	0.2	*	0.77	1.06	-1.1	*	-1.1	*	-2.2
El Salvador	Median	0.2	-0.5	*	0.88	1.26	-3.2	*	4.8	*	1.6
Ghana	$1.25/day	1.5	-0.8		0.82	1.03	-1.0	*	-1.4		-2.4
Kenya	$1.25/day	1.0	0.0	*	0.85	1.23	-1.1	*	-0.5		-1.6
Malawi	$1.25/day	0.0	0.3		0.85	0.94	-4.3	*	3.0	*	-1.3
Mexico	Median	-0.8	0.4		1.03	1.61	0.4		-0.3		0.1
Nicaragua	Median	0.2	-0.6		0.77	1.13	-0.8		0.4		-0.4
Nigeria	$1.25/day	-0.8	0.7		1.26	1.13	-0.8		0.4		-0.4
Peru	Median	-4.7	3.3		1.05	0.77	0.1		-0.2		-0.1
Rwanda	$1.25/day	-1.0	0.0	*	0.77	1.00	-0.3		0.1		-0.2

Bias is in terms of percentage points. Differences in targeting inclusion and exclusion are also in terms of percentage points.
For bias, an asterisk means that the scorecard's or the PAT's bias lies outside the other's 90-percent confidence interval.
For targeting measures, an asterisk means that the scorecard's figure is outside the 90-percent confidence interval of the PAT's

Table 4: For each year in which USAID's *Microenterprise Results Reporting* includes PAT estimates, the number of countries for which a PAT exists, the number of USAID microenterprise partners who report estimates based on a PAT, and the estimated poverty rate by the USAID "very poor" line of the participants of reporting partners

Fiscal Year	PATs	Partners reporting	% "Very poor"
2007	17	31	22
2008	26	18	29
2009	29	16	23
2010	33	10	Not reported
2011	38	10	38
2012	38	4	56

The count of PATs includes two scorecards (Haiti and Philippines) branded as PATs.

Source: "Microenterprise Results Reporting" reports by USAID to the U.S. Congress, various years.

Table A1: Household- and person-level poverty rates for the scorecard and the PAT when applied to national-survey data for the median and $1.25/day poverty lines, and whether an uncorrected scorecard or PAT correctly implements the USAID "very poor" line

	Rate by median line (%)				Rate by $1.25/day (%)				"Very poor" line		
	HH level		Person level		HH level		Person level			Correct?	
Country	Card	PAT	Card	PAT	Card	PAT	Card	PAT	Line	Card	PAT
Cambodia	14.4	15.3	17.4	—	33.2	35.2	37.9	40.1	$1.25/day	Yes	No
Ecuador	15.4	14.8	19.7	—	—	2.1	2.4	—	Median	No	No
El Salvador	19.2	20.2	23.2	—	46.2	5.4	52.7	—	Median	Yes	No
Ethiopia	—	—	—	—	27.2	31.5	35.9	39.8	$1.25/day	No	No
Ghana	8.9	14.2	14.2	—	16.2	18.9	25.3	28.8	$1.25/day	Yes	No
Guatemala	20.5	23.0	28.1	—	3.6	—	5.4	—	Median	Yes	No
Kenya	17.7	19.1	23.3	—	36.3	36.3	46.1	—	$1.25/day	No	Yes
Malawi	19.6	21.9	26.2	—	66.1	64.8	74.2	73.1	$1.25/day	No	No
Mexico	19.2	20.1	23.7	—	1.6	1.8	2.2	—	Median	Yes	No
Nepal	12.3	14.9	15.3	—	47.8	52.6	53.6	—	$1.25/day	Yes	No
Nicaragua	17.5	19.6	23.8	—	8.3	8.4	11.3	—	Median	Yes	No
Nigeria	21.0	22.6	27.3	—	58.0	58.2	68.5	—	$1.25/day	—	No
Peru	—	15.9	17.4	—	0.5	—	0.6	—	Median	—	No
Rwanda	26.1	26.9	28.3	—	69.5	72.9	71.7	—	$1.25/day	Yes	No
Tanzania	12.6	—	16.8	—	61.3	43.0	71.3	—	$1.25/day	Yes	No

'—' means that a rate was not reported and could not be inferred.

Poverty rates for the scorecard are as originally documented or as corrected (if originally in error).

Table A2: Household- and person-level poverty rates for the scorecard and the PAT when applied to national-survey data, and whether a (uncorrected) tool correctly implements the national poverty line

	Rate by national line (%)				Exp., deflators, calcs., and natl. line correct?	
	HH level		Person level			
Country	Card	PAT	Card	Official	Card	PAT
Cambodia	30.2	30.6	34.7	34.7	Yes	No
Ecuador	30.9	29.7	39.4	38.3	No	Yes
El Salvador	39.9	40.4	46.4	—	Yes	Yes
Ethiopia	—	—	—	—	—	—
Ghana	18.9	28.4	28.5	28.5	Yes	No
Guatemala	45.8	46.0	56.1	56.2	Yes	Yes
Kenya	38.3	38.3	46.6	46.6	Yes	Yes
Malawi	43.6	43.8	52.4	52.4	Yes	Yes
Mexico	40.7	40.2	47.4	47.4	Yes	No
Nepal	25.9	29.8	30.8	30.9	Yes	No
Nicaragua	39.2	39.2	48.4	48.3	Yes	Yes
Nigeria	44.8	45.2	54.5	54.4	Yes	Yes
Peru	29.0	31.8	34.8	34.8	Yes	No
Rwanda	54.1	53.9	56.9	56.9	Yes	Yes
Tanzania	26.6	—	33.6	33.6	Yes	No

'—' means that a rate was not reported and could not be inferred.

Poverty rates are as originally documented.

Table A3: Person-level poverty rates for the $1.25/day poverty line for the scorecard and PovcalNet, and which estimate is to be preferred based on the country-level notes in the appendix

	Rate by $1.25/day (%)		
	Person level		
Country	Card	PovCalNet	Which is to be preferred?
Cambodia	37.9	37.7	Either
Ecuador	2.4	9.1	Scorecard
El Salvador	52.7	11.0	Scorecard
Ethiopia	35.9	39.0	Scorecard
Ghana	25.3	28.6	Scorecard
Guatemala	5.4	11.9	Scorecard
Kenya	46.1	43.4	Scorecard
Malawi	74.2	73.9	Either
Mexico	2.2	1.2	Scorecard
Nepal	53.6	53.1	Scorecard
Nicaragua	11.3	11.9	Scorecard
Nigeria	68.5	63.1	Scorecard
Peru	0.6	5.5	Scorecard
Rwanda	71.7	72.1	Either
Tanzania	71.3	67.9	Scorecard

Poverty rates for the scorecard are as originally documented or as corrected (if originally in error).

PovcalNet figures were retrieved 15 January 2014 from iresearch.worldbank.org/PovcalNet/index.htm.

For justification of indication of preference, please see detailed country-level notes.