

SUPPLEMENTARY MATERIALS

1-2-3 SURVEYS

As suggested by its name, the *1-2-3 Survey* is a three-phase survey. The basic rationale of this tool is the following. The first phase is an augmented labour force survey (LFS). It documents and analyses labour market functioning and is used as a filter for the second phase to identify a representative sample of IF heads who are then interviewed. Phase 2 is designed to measure the firms' main economic and productive characteristics (production, value-added, investment and financing), the main difficulties encountered developing the business, and demand for public support by informal entrepreneurs. Lastly, in the third phase, a specific income/expenditure survey is administered to a sub-sample of households selected from phase 1 to estimate the weights of the formal and informal sectors in household consumption by product and household type. Phase 3, not used in this paper, is also used to estimate households' living standards and monetary poverty, based on either income or expenditure.

The following presents a brief description of the sampling plan and the content of the questionnaires administered. The sampling frame for the LFS (phase 1) was a classic two-stage stratified sample. The primary sampling units were census enumeration areas (EAs). In each EA (125 per country), some 20 households were randomly selected with equal probabilities (see Brilleau, Roubaud and Torelli, 2005a for more details). The final sample size was 18,000 households (approximately 2,500 households in each country, except in Benin where 3,000 households were interviewed). The IFs were stratified for phase 2 using phase 1's rich data. Twenty strata were defined by industry (10 industries) and IF head status (employer and/or own account worker). The unequal probabilities for each stratum were calculated based on the number of IFs and their economic potential in terms of development policies (employer and industrial IFs are overrepresented). Here again, IFs were randomly selected from phase 1 master samples for each stratum. Lastly, 6,500 IFs completed the questionnaires (around 1,000 IFs per country). Sample weights were computed by adjusting the theoretical probability of inclusion by the usual post-stratification procedures.

Note here that the *1-2-3 surveys* have three major advantages over other alternative datasets. First, the mixed household-enterprise survey frame is the only way to ensure the full representativeness of the informal sector (ILO, 2013). Previous studies, especially in SSA, are typically based on enterprise surveys, which cover just part of all informal firms; the upper tier of the informal sector generally works in specific industries, mainly manufacturing. Second, our seven surveys are fully comparable, as sampling methodology and questionnaire were kept constant across all countries. To our knowledge, this is the first time ever that such a series of repeated cross sections on representative IFs has been made available for SSA. Third, given that IFs do not usually keep books (and, in many cases, do not even have any written records), the survey questionnaire was designed to help the owners of the firms establish all their sales and expenses product by product over a flexible period of time adjusted to each individual case. This detailed and comprehensive information collection procedure is the only way to get reliable data and avoid the usual underestimation biases caused by more aggregate questions. The same extensive process was used to reconstitute the stock of capital evaluated at replacement cost to take into account depreciation.

In the *1-2-3 surveys*, the criteria used to identify IFs are the absence of an administrative registration number and/or of written bookkeeping, excluding farming activities. An aggregate three-industry nomenclature shows that trade accounts for a large share of the informal sector (Table 1). A full 46 per cent of IFs work in trade, as opposed to 28 per cent in manufacturing (including construction) and 26 per cent in services. The dominance of trade is observed in almost all the capital cities. Its share ranges from 40 per cent in Abidjan to 52 per cent in Bamako. One of the upshots of these

results is that the usual surveys covering solely manufacturing overlook the majority of informal firms.

Table S1: Informal firms' distribution by industry (%)

| | Cotonou | Ouagadougou | Abidjan | Bamako | Niamey | Dakar | Lomé | Total |
|--------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Manufacturing | 21.9 | 34.2 | 28.5 | 27.3 | 43.2 | 31.1 | 23.0 | 28.4 |
| Clothing, leather, shoes | 9.2 | 7.5 | 12.4 | 10.9 | 8.2 | 7.6 | 9.1 | 10.1 |
| Other manufacturing | 8.1 | 21.1 | 9.4 | 10.3 | 32.0 | 15.9 | 10.2 | 12.4 |
| Construction | 4.6 | 5.6 | 6.7 | 6.2 | 3.0 | 7.6 | 3.8 | 5.9 |
| Trade | 49.2 | 48.7 | 40.0 | 51.5 | 40.6 | 47.3 | 48.5 | 45.5 |
| In-shop retail and | 13.5 | 11.4 | 11.1 | 9.1 | 7.3 | 11.1 | 11.9 | 11.1 |
| Out-of-shop retail | 35.7 | 37.3 | 28.9 | 42.4 | 33.3 | 36.2 | 36.5 | 34.4 |
| Services | 28.9 | 17.1 | 31.5 | 21.3 | 16.2 | 21.6 | 28.5 | 26.1 |
| Catering | 10.5 | 4.8 | 7.0 | 3.0 | 0.5 | 4.1 | 7.0 | 6.0 |
| Repairs | 3.5 | 4.8 | 6.0 | 2.7 | 2.8 | 2.1 | 5.3 | 4.3 |
| Transport | 5.2 | 1.0 | 4.1 | 2.9 | 1.9 | 4.3 | 4.4 | 3.8 |
| Other services | 9.7 | 6.4 | 14.4 | 12.7 | 10.9 | 11.1 | 11.8 | 12.0 |
| Total | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Source: Authors' calculations based on the 1-2-3 surveys, phase 2.

Table S2: City and sector fixed effect estimates in regressions on the probability of paying a bribe

| | (1) | (2) | (3) | (4) | (5) | (6) |
|--|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| Country fixed effect (ref. Togo) | | | | | | |
| Benin | -0.67*** (0.16) | -0.57*** (0.17) | -0.66*** (0.17) | -0.63*** (0.16) | -0.62*** (0.16) | -0.22** (0.10) |
| Burkina Faso | -0.73*** (0.18) | -0.76*** (0.20) | -0.77*** (0.19) | -0.77*** (0.19) | -0.70*** (0.19) | -0.06 (0.11) |
| Ivory Coast | -0.08 (0.14) | -0.03 (0.14) | -0.13 (0.15) | -0.12 (0.14) | -0.06 (0.14) | -0.08 (0.10) |
| Mali | -0.25* (0.15) | -0.20 (0.15) | -0.31** (0.15) | -0.28* (0.15) | -0.21 (0.16) | -0.17 (0.11) |
| Niger | -0.28 (0.17) | -0.35* (0.19) | -0.24 (0.17) | -0.23 (0.17) | -0.21 (0.17) | 0.08 (0.12) |
| Senegal | -0.08 (0.14) | -0.03 (0.13) | -0.08 (0.14) | -0.08 (0.14) | -0.08 (0.14) | 0.18* (0.10) |
| Sector fixed effects (ref. Transport) | | | | | | |
| Clothing, leather, shoes | -0.71*** (0.16) | -0.72*** (0.16) | -0.60*** (0.16) | -0.57*** (0.16) | -0.49*** (0.16) | -0.65*** (0.12) |
| Other manufacturing | -0.89*** (0.17) | -0.87*** (0.17) | -0.81*** (0.17) | -0.79*** (0.17) | -0.79*** (0.17) | -0.68*** (0.11) |
| Construction | -1.11*** (0.23) | -1.25*** (0.23) | -1.04*** (0.25) | -1.01*** (0.25) | -0.99*** (0.24) | -1.00*** (0.17) |
| In-shop retail and wholesale | -0.58*** (0.15) | -0.57*** (0.16) | -0.56*** (0.16) | -0.51*** (0.15) | -0.55*** (0.16) | -0.58*** (0.11) |
| Out-of-shop retail | -0.63*** (0.15) | -0.52*** (0.17) | -0.49*** (0.16) | -0.45*** (0.16) | -0.50*** (0.16) | -0.49*** (0.11) |
| Catering | -1.11*** | -0.90*** | -0.82*** | -0.81*** | -0.82*** | -0.72*** |

| | | | | | | |
|----------------|----------|----------|----------|----------|----------|----------|
| Repairs | (0.22) | (0.24) | (0.23) | (0.23) | (0.26) | (0.14) |
| | -0.57*** | -0.75*** | -0.67*** | -0.60*** | -0.59*** | -0.61*** |
| Other services | (0.18) | (0.18) | (0.19) | (0.18) | (0.18) | (0.13) |
| | -1.20*** | -1.23*** | -1.12*** | -1.10*** | -1.07*** | -0.76*** |
| | (0.21) | (0.22) | (0.21) | (0.21) | (0.22) | (0.13) |

Note: The estimated models are those reported in Table 3.

Standard errors are clustered at neighbourhood level. ***, ** and * denote significance at the 1%, 5% and 10% level respectively.

Table S3: City and sector fixed effect estimates in regressions on the bribe amount paid

| | (1) | (2) | (3) | (4) |
|--|----------|----------|----------|----------|
| Country fixed effect (ref. Togo) | | | | |
| Benin | -4.60*** | -3.04*** | -2.82*** | -2.73*** |
| | (0.78) | (0.64) | (0.69) | (0.68) |
| Burkina Faso | -4.69*** | -3.41*** | -4.94*** | -3.31*** |
| | (0.79) | (0.74) | (0.62) | (0.62) |
| Ivory Coast | -2.17*** | -0.90 | -1.46*** | -0.77 |
| | (0.72) | (0.59) | (0.52) | (0.55) |
| Mali | -1.89** | -0.21 | -0.91* | -0.52 |
| | (0.81) | (0.60) | (0.53) | (0.52) |
| Niger | -2.48*** | -1.56** | -1.87*** | -1.73*** |
| | (0.79) | (0.64) | (0.56) | (0.59) |
| Senegal | -1.62* | -0.25 | -0.06 | 0.11 |
| | (0.90) | (0.57) | (0.59) | (0.62) |
| Sector fixed effects (ref. Transport) | | | | |
| Clothing, leather, shoes | -3.08*** | -4.05*** | -2.55*** | -2.28*** |
| | (0.69) | (0.81) | (0.80) | (0.75) |
| Other manufacturing | -3.31*** | -5.14*** | -4.07*** | -3.44*** |
| | (0.70) | (0.73) | (0.66) | (0.66) |
| Construction | -4.59*** | -7.51*** | -5.36*** | -5.38*** |
| | (0.80) | (0.77) | (0.81) | (0.75) |
| In-shop retail and wholesale | -0.70 | -1.15 | -2.72*** | -1.42* |
| | (0.74) | (0.81) | (0.77) | (0.73) |
| Out-of-shop retail | -0.77 | -1.21* | -2.15*** | -1.47** |
| | (0.61) | (0.66) | (0.67) | (0.66) |
| Catering | -4.91*** | -5.22*** | -3.66*** | -3.19*** |
| | (0.80) | (0.81) | (0.80) | (0.81) |
| Repairs | -1.31 | -3.41*** | -2.61*** | -2.05*** |
| | (0.88) | (0.80) | (0.87) | (0.77) |
| Other services | -4.26*** | -6.42*** | -5.01*** | -4.36*** |
| | (0.79) | (0.88) | (0.79) | (0.82) |

Note: The estimated models are those reported in Table 4.

Standard errors are clustered at neighbourhood level. ***, ** and * denote significance at the 1%, 5% and 10% level respectively.

Table S4: Bribery and Informal firms' turnover: Average Treatment Effects (ATEs)

| | Coef. | Std errors | # of obs. |
|--|-------|------------|-----------|
|--|-------|------------|-----------|

| Panel A: Inverse Probability Weighting Model (IPW) | | | |
|---|---------|------|-------|
| Full sample | 0.27 | 0.27 | 6,442 |
| <i>Top performers</i> | -0.14 | 0.56 | 645 |
| <i>Constraint Gazelles</i> | 0.31*** | 0.08 | 2,973 |
| <i>Survivalists</i> | 0.58 | 0.39 | 2,824 |
| Panel B: Propensity Score Matching Model (PSM) | | | |
| Full sample | 0.40 | 0.15 | 6,442 |
| Top performers | -0.06 | 0.39 | 645 |
| Constraint Gazelles | 0.28** | 0.14 | 2,973 |
| Survivalists | 0.23 | 1.14 | 2,824 |

Note: In panel A and B, the incidence of corruption is explained by sales per employee (in log), the capital/labour ratio (in log), dummies denoting IF size and the fact that the entrepreneur pays taxes as well as 3-digit sector indicators and city fixed effects. Figures in parentheses are probability tests. Standard errors are clustered at neighbourhood level. ***, ** and * denote significance at the 1%, 5% and 10% level respectively.