



HAL
open science

Government Effectiveness and Regional Variation in Informal Employment

Erik Jonasson

► **To cite this version:**

Erik Jonasson. Government Effectiveness and Regional Variation in Informal Employment. The Journal of Development Studies, 2012, 48 (04), pp.481-497. 10.1080/00220388.2011.615922. hal-00815560

HAL Id: hal-00815560

<https://hal.science/hal-00815560>

Submitted on 19 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Government Effectiveness and Regional Variation in Informal Employment

Journal:	<i>Journal of Development Studies</i>
Manuscript ID:	FJDS-2010-Sep-0040.R2
Manuscript Type:	Original Manuscripts
Keywords:	informal employment, government effectiveness, Latin America < Geographical Area, Brazil

SCHOLARONE™
Manuscripts

Government Effectiveness and Regional Variation in Informal Employment

This paper analyses the role of government effectiveness in the determination of informal employment. A theoretical model is developed, in which local governance and worker skill level are assumed to influence the decision of the worker whether to seek employment in the formal or informal sector. The model is assessed empirically using data from Brazil, where almost half of the urban labour force is employed informally. The empirical analysis supports the predictions of the model and suggests that the probability of a worker being employed informally is lower in regions with better governance and higher average education.

1. Introduction

Widespread informal employment continues to be a salient feature of labour markets in developing countries. In their comprehensive study of informality in Latin America, Perry et al. (2007) estimate the share of the labour force employed informally to range from about 20 per cent in Chile to about 80 per cent in Paraguay. Despite economic growth, recent evidence suggests that informal employment has not declined in developing countries over the past two decades (Freeman, 2009).

The defining characteristic of informal employment is usually that it is not regulated in any formal employment contract, and therefore is not subject to labour market regulations. To the worker, this implies being outside most social security arrangements while faced often with low compensation

1
2
3 and poor working conditions (Jütting et al., 2008). For businesses, being informal constrains the
4 ability to raise financial capital and to enforce complex contracts, which is likely to restrict their
5 prospects of growing (Loayza, 1996; Fajnzylber et al, 2009; Hansen et al, 2009). To the economy as a
6 whole, wide-spread informality means difficulties in collecting tax revenues for the provision of
7 public goods and services. Recent empirical evidence even suggests that informality affects economic
8 growth negatively, controlling for other country characteristics (Loayza et al., 2010).
9

10
11
12
13
14
15
16 A relatively large theoretical and empirical literature has sought to identify the determinants of
17 informality, both in terms of worker and firm characteristics, on the one hand, and the burden of
18 formal rules and regulations on the other hand. Empirical evidence also suggests that the manner in
19 which governments choose to enforce regulations is important for the degree of informality in a
20 country. Friedman et al. (2000), for example, in a study including 67 countries, suggest that high taxes
21 per se do not drive businesses into the informal sector. Instead, it is the discretion in the application of
22 rules and the corresponding corruption that it generates that seems to make businesses escape the
23 formal economy.
24
25
26
27
28
29
30
31
32

33
34 This paper adds to this branch of literature by assessing the role of government effectiveness in
35 determining the probability of informal employment among Brazilian workers. Government
36 effectiveness is understood here as the extent to which the government supports and implements
37 formal institutions and provides high-quality public goods and services. It corresponds closely to the
38 definition given by Kaufmann et al. (2009). Jonasson (2011) reports a negative correlation between
39 government effectiveness and the size of the informal sector among Brazilian municipalities. This
40 paper extends on those findings by developing a theoretical model, which incorporates government
41 effectiveness in the worker's decision of whether to participate in the formal or informal sector. The
42 model is evaluated empirically using Brazilian worker-level data. Among Brazil's over 5500
43 municipalities, informality varies between 20 and 80 per cent when measured as the share of the urban
44 labour force lacking a signed labour card. This wide range of informality among municipalities
45 suggests that labour and tax regulations, which are determined largely at the national level, are far
46 from the only determinants of informality in a country. By using municipal data on public sector
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 management and a range of other indicators, and keeping country-specific characteristics fixed, the
4 effects of government effectiveness are assessed.
5
6
7
8

9 10 **2. Related literature**

11
12 This paper relates primarily to the literature on institutional determinants of informal economic
13 activity. Institutional determinants refer here both to formal institutions, such as labour and tax
14 regulations, and informal institutions, such as social norms of tax compliance (Casson et al., 2010).
15 Informal economic activities include any market-based legal production of goods and services,
16 deliberately hidden from authorities to avoid taxation or costly regulation, and also working, or hiring
17 labour, without an employment contract complying with labour market regulations. There is a large
18 literature on the determinants of informality (reviewed by Kucera and Xenogiani, 2009; Perry et al.,
19 2007; and Schneider and Enste, 2000) and a growing literature on how various notions of governance
20 and social norms shape economic development and other aspects of society. Most such studies,
21 however, focus on economic growth as the outcome of interest (e.g. Glaeser et al., 2004; Hall and
22 Jones, 1999). Relatively few studies from these two branches of literature are concerned specifically
23 with the role of government effectiveness in shaping the informal sector. The empirical evidence that
24 exists is to a large extent at cross-country level and little work has been done at the sub-national or
25 micro level.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

42 Empirical cross-country evidence by Johnson et al. (1998) and Friedman et al. (2000) suggests
43 that the unofficial economy tends to be larger, not where taxes are higher, but in countries where the
44 regulatory burden is higher, corruption is more wide-spread, and where the rule of law is weaker.
45 Loayza et al. (2006), for a sample of about 70 countries, find that the level of business regulation is on
46 its own correlated positively with informality but when quality of governance is interacted with the
47 level of regulation, regulation is negatively related with informality. Loayza et al. (2010) analyze a
48 broader set of determinants of informality for a sample of about 85 countries and find that informality
49 is negatively related with law and order, business regulatory freedom, and with schooling
50 achievement. Restricting the focus to Latin America, “policy and institutional variables related to the
51 quality of the state are the most important factors explaining the differences in informality” (Loayza et
52
53
54
55
56
57
58
59
60

1
2
3 al., 2010, p. 179). Torgler and Schneider (2009) add tax morale to the set of explanatory factors for
4
5 cross-country variation in informality; various indicators of good governance relate negatively to the
6
7 size of the shadow economy, but also the willingness of citizens to pay taxes is related negatively with
8
9 the size of the shadow economy. While the use of various indicators of quality of governance has
10
11 been questioned (Langbein and Knack, 2010), it is probably fair to conclude from the above studies
12
13 that it is not only the design of formal rules and regulations that determines the size of the informal
14
15 sector, but also the manner in which they are implemented and obeyed.
16
17

18
19 At sub-national level, there are a few empirical studies concerned with regional variation in the
20
21 size of the informal sector, e.g. Chaudhuri et al. (2006) on 14 states in India and Torgler and
22
23 Schneider (2007) on the 26 cantons of Switzerland. At the micro level, few empirical studies analyse
24
25 the effect of institutional factors on the propensity of workers or businesses to participate in the
26
27 informal sector. A recent exception is the study by Almeida and Carneiro (2009), who analyse how
28
29 differences in enforcement of labour regulation across regions in Brazil affect regional informal
30
31 employment and unemployment. They find evidence that stricter enforcement in a region, measured
32
33 as the aggregate amount of registration fines issued, leads to less informal employment but also to
34
35 higher unemployment. Empirical studies at the micro level otherwise tend to be concerned primarily
36
37 with worker characteristics as explanatory factors of informal employment. These studies generally
38
39 conclude that the probability of working in the informal sector decreases with human capital
40
41 endowment measured by years of education, job tenure, and experience (Funkhouser, 1996; Gong et
42
43 al., 2004; Marcouiller et al., 1997).
44
45
46

47 In sum, empirical evidence at the cross-country level suggests that corruption and low quality
48
49 of governance, as well as and low tax morale, seem to cause informality to a greater extent than strict
50
51 labour regulation and high taxes. Empirical evidence at the micro level provides strong evidence that
52
53 worker-specific characteristics fundamentally affect the probability of working in the formal or
54
55 informal sector. Less is known about the causes of regional variation in informality within countries.
56
57 The fact that it varies considerably across regions suggests that region-specific characteristics could
58
59 be at least, or even more, important determinants of informality than labour and tax regulations and
60
other nation-wide formal institutions.

3. Theoretical framework

In this section, a static model is developed to explain regional variation in informal employment. The proposed determinants of such variation are regional differences in worker skill distribution, tax rates and government effectiveness. The model is inspired by Loayza and Rigolini (2011), Galiani and Weinschelbaum (2010), and to a lesser extent by Rauch (1991), the principal novelty of the model being the inclusion of government effectiveness as a factor influencing the worker's decision on sector of employment. The basis of the model is a two-sector framework of occupational choice, in which the worker chooses sector depending on expected earnings. The two sectors of employment here are the formal and the informal sector.

Consider first the case in which the worker only takes into account expected income when choosing sector of employment. Utility is assumed to increase monotonically in income, so that the worker maximizes expected utility by choosing to work in the sector that gives the highest expected income. The decision to work is pre-determined outside the model; hence the consumption/leisure decision is abstracted from. The utility maximisation problem reduces to:

$$\max_j U(y^j), \quad j = (FS, IS) \quad (1)$$

where y^j is expected labour income in sector j . Workers are assumed to differ by skill endowment s , which is distributed among workers according to some density function $h(s)$ ($0 \leq s \leq 1$). Skill endowment determines the productivity of the worker, measured in number of homogenous labour units, $L(s)$. The formal sector (FS) and the informal sector (IS) are assumed to differ from each other in several aspects, affecting the expected income and hence the incentives faced by the worker in the choice of sector.

First, the institutional arrangements in the formal sector are assumed to allow for higher labour productivity than in the informal sector. The better the local government serves its citizens (i.e. the higher the level of government effectiveness), the wider is the labour productivity gap between

sectors. The model allows for different channels through which this takes place. A broad interpretation is that, with formal institutions that are well implemented and enforced by the government, firms face better prospects for benefiting from economies of scale through cheaper access to capital and better ability to enforce complex contracts (de Paula and Scheinkman, 2010; Straub, 2005). As a consequence, production in the formal sector will be more capital intensive on average, hence increasing the productivity of labour. A narrower interpretation is that local governments can supply public services and support institutions that directly affect labour productivity, independently of production technology, in an efficiency-wage manner. It could be that the worker is more productive in the formal sector due to an increased sense of transparency, security, and stability in the work environment of the formal sector, compared to the informal sector.

Let labour productivity be characterized by positive but diminishing returns to worker skill in both sectors. In the informal sector, labour productivity, measured in units of units of homogenous labour (L), is given by:

$$L^I(s) = s^\alpha \quad (0 < \alpha < 1) \quad (2)$$

If g denotes the local level of government effectiveness, then labour productivity in the formal sector is given by:

$$L^F(s, g) = (1 + g)s^\alpha \quad (-1 < g < 1) \quad (3)$$

Second, assume that income is taxed at local rate t in the formal sector. While workers in the informal sector do not pay taxes, they face a risk of being caught by the authorities for working informally (or for their employers hiring them informally). In case of detection, labour income will be zero. The risk of being caught, c , is determined by the strength of the enforcement of tax and labour regulation e ($0 < c < 1$; $e > 0$; $c'(e) > 0$). Expected income in the informal sector is given by:

$$y^I = [1 - c(e)]s^\alpha \quad (4)$$

Third, expected income in the formal sector is affected by the worker's probability of finding a job in that sector and thereby gaining the corresponding productivity enhancement. The worker's perceived probability of obtaining a job is an increasing function of skills, given by:¹

$$\pi(s) = s^\beta \quad (0 < \beta < 1) \quad (5)$$

The formal and the informal sectors are assumed to be integrated in the sense that one unit of homogenous labour, L , is paid the same in both sectors. For simplicity, assume that the worker faces a perfectly elastic labour demand at a price equal to one.² Imposing labour income tax rate t ($0 < t < 1$), the expected net income in the formal sector is:

$$y^{FS} = (1 - t)(1 + g)s^{\alpha + \beta} \quad (6)$$

The threshold level of skill that equates the expected incomes y^{IS} and y^{FS} in the two sectors is:

$$\underline{s} = \left[\frac{1 - c}{(1 - t)(1 + g)} \right]^{\frac{1}{\beta}} \quad (7)$$

Whenever the skill level falls short of \underline{s} the individual does not expect to benefit from participating in the formal sector. Given that workers choose the sector rationally (and do not attach utility or disutility to any of the sectors beyond labour income), $y^{IS}(\underline{s}) = y^{FS}(\underline{s})$ denote the highest expected earnings in the informal sector and the lowest in the formal sector, respectively, for given levels of t , g and c . The existence of a formal sector requires that \underline{s} is less than 1, or:

$$\gamma \equiv \frac{(1 - t)(1 + g)}{(1 - c)} > 1 \quad (8)$$

where γ denotes the ratio of formal to informal income for the highest educated (for whom $s = 1$).

The higher the skill threshold, the higher is the probability that any given worker will prefer to work in the informal sector. Differentiating (7) with respect to t , g , and c shows that the propensity to work informally increases with t and decreases with g and c . Figure 1 provides an illustration of these

effects. Below the skill threshold \underline{s}_0 the worker expects to earn more by being in the informal sector than in the formal sector, depicted by the segment of the y_0^{IS} curve above the y_0^{FS} line. The reason for this is the low perceived probability for low-skilled workers to find a formal job.³ The effect of an increase in law enforcement, which increases c , is shown by the lowering of the informal-sector income curve from y_0^{IS} to y_1^{IS} , which decreases the skill threshold from \underline{s}_0 to \underline{s}_1 . On the other hand, an increase in the tax rate or a decline in government effectiveness lowers the formal-sector income curve from y_0^{FS} to y_1^{FS} , increasing the skill threshold from \underline{s}_0 to \underline{s}_2 .

[Figure 1]

The share of the labour force that works in the informal sector in region m is given by:

$$IS_m = H_m[\underline{s}_m(g_m, t_m, e_m)] = \int_0^{\underline{s}_m} h_m(s) ds \quad (9)$$

where $H_m(s)$ is the cumulative density function of skill endowments s in regional m . Thus, the model predicts that regional differences in the relative size of the informal sector are determined by regional differences in the quality of institutions g , the tax rate t , the risk of income loss in the informal sector c , and in the skill distribution in the local labour force, $h_m(s)$. Moreover, given the levels of t , g , and c , informality will be higher (lower) in regions with lower (higher) average level of skill.

The model described thus far does not predict that there could be an overlap between the sectors in skills and/or earnings, i.e. that some workers in the informal sector are more skilled and earn more than some workers in the formal sector. Such an overlap has been found to be extensive in several empirical studies (Bosch et al., 2007; Maloney, 2004) and is evident also in the case of Brazil (Table 1, below). One explanation for such a sectoral overlap is the existence of worker-specific non-pecuniary costs and benefits of participation in the informal sector. Non-pecuniary benefits of working in the informal sector could be a higher degree of flexibility in working hours or a greater sense of freedom and independence (Maloney, 2004; Marcouiller, et al., 1997). A perceived cost of informal work could be the sense of insecurity about one's livelihood, in terms of future earnings,

employment contract renewal, or enterprise survival (Jütting, et al., 2008). Another cost could be the moral, or psychic, cost of violating tax or labour regulations (Torgler and Schneider, 2009).

To take non-pecuniary costs of informal employment and benefits into account, the utility maximisation problem in (1) is adjusted slightly to:

$$\max_j U(\hat{y}^j), \quad j = (FS, IS) \quad (1')$$

where \hat{y}^j denotes total expected benefit of employment in sector j . Let $b(x)$ denote the net value of the non-pecuniary benefits and costs for the worker of being in the informal sector ($-1 < b < 1$), x denoting worker characteristics affecting the valuation of these net benefits. Assume that b is proportional to income such that the sum of benefits in the informal sector is:

$$\hat{y}^{IS} = [1 + b(x)]y^{IS} = [1 + b(x)] \times [1 - c(e)]s^\alpha \quad (10)$$

The worker skill threshold equating total benefits in the informal sector with income in the formal sector, \hat{s} , is given by:

$$\hat{s} = \left[\frac{(1-c)(1+b)}{(1-t)(1+g)} \right]^{\frac{1}{\alpha}} \quad (11)$$

The qualitative effects of changes in t , c , and g are the same as in the case without non-pecuniary benefits. The additional effect to note is the positive relationship between the threshold level \hat{s} and the net benefits b .

Figure 2 illustrates situations in which the worker experiences zero, positive, and negative net benefits in the informal sector. For given levels of t , c , and g , let \hat{s}_0 be the skill threshold in the case when the net benefit b is zero. If a worker values the net benefit by $b_1 > 0$, then the total benefit in the informal sector is given by the \hat{y}^{IS}_1 curve, which lies above the y^{IS} curve. The \hat{y}^{IS}_1 curve intersects the formal sector income curve, y^{FS} , at $\hat{s}_1 > \hat{s}_0$. The worker's monetary income in the informal sector at \hat{s}_1 , $y^{IS}(\hat{s}_1)$, is lower than the income she expects to earn in the formal sector with the same skill level, $y^{FS}(\hat{s}_1)$. Due to the non-pecuniary benefit, however, she is indifferent between the two sectors at this

1
2
3 skill threshold. If, instead, a worker experiences a negative net benefit $b_2 < 0$ of participating in the
4 informal sector, then the total benefit curve falls below y^{IS} and the skill threshold decreases from \hat{s}_0 to
5 \hat{s}_2 . Despite the fact that $y^{IS}(\hat{s}_2)$ is higher than $y^{FS}(\hat{s}_2)$, the worker is indifferent between the two sectors
6
7
8 at this point, due to the disutility attached to work in the informal sector.
9
10
11

12
13
14 [Figure 2]
15

16
17
18 Given that b is individual specific, there will be an entire distribution of thresholds \hat{s} within
19 each region. Let $s^* = s - \hat{s}$, so that whenever s^* is below 0, the individual works in the informal sector.
20
21 If the distribution of s^* is given by $k(s^*)$, the size of the informal sector in region m is given by:
22
23

$$24 \quad IS_m[k_m(s^*)] = \int_{-1}^0 [k_m(s^*)] ds^* \quad (12)$$

25
26
27
28
29
30 The empirical analysis in the remainder of the paper is concerned with testing the hypothesis of
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

4. Empirical approach

41
42
43 The theoretical model is evaluated using worker-level and municipal data from Brazil. A
44 binomial probability model is estimated at worker level to assess the extent to which the individual
45 and local factors considered in the model correlate with the probability of having informal
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

41
42
43 The binomial probability model is estimated using a probit model. The difference s^* between
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
45

s_i^* is below zero. Thus, if s_i^* is determined by the set of exogenous variables under consideration, then the probability that the individual participates in the informal sector is given by:

$$p_{im} \equiv \text{prob}(is_{im} = 1) = \text{prob}(s_{im}^* < 0 | s_{im}, x_{im}, g_m, z_m) \quad (13)$$

The binary variable is_{im} equals one if the individual works in the informal sector and zero if in the formal sector (subscript m denoting municipality). s_{im} is a set of proxies for worker skills; x_{im} is a vector of other individual characteristics that are assumed to affect productivity and the individual-specific valuation of the non-pecuniary net benefit in the informal sector; g_m is a set of proxies for government effectiveness; and z_m , finally, consists of local productivity shifters that might affect the relative productivity in the formal and informal sectors. The probability model is estimated under the assumption that s_{im}^* is a linear function of the exogenous variables:

$$s_{im}^* = s_{im}\beta_s + x_{im}\beta_x + g_m\beta_g + z_m\beta_z + u_{im} \quad (14)$$

where the β s denote vectors of parameters to be estimated and u_{im} denotes a residual that is assumed to be normally distributed with zero mean and variance σ^2 . Let F_u denote the cumulative distribution function of u . The probit model then gives:

$$\hat{p}_{im} = 1 - F_u(s_{im}\hat{\beta}_s + x_{im}\hat{\beta}_x + g_m\hat{\beta}_g + z_m\hat{\beta}_z) \quad (15)$$

where \hat{p}_{im} denotes the estimated p_{im} and the $\hat{\beta}$ s are the estimated coefficients ($0 \leq \hat{p} \leq 1$).

If the variable representing government effectiveness is correlated with the error term u_{im} in (14), the coefficient estimate will be biased. Such correlation could be due to unobserved regional factors, such as culture or aspects of economic development not captured in z_m , that affect government effectiveness while also affecting the probability that a worker is employed informally. In the IV probit approach, a set of instrumental variable candidates are used in order to remove the potential bias in the coefficient estimate for government effectiveness. The model above is extended as follows. To simplify the notation, let X denote the set of all exogenous variables in (14), excluding government

effectiveness, GE , which is now assumed to be endogenous. Let β denote the set of all corresponding coefficients. Finally, let Z denote the set of instrumental variables. The IV probit model estimates:

$$s_{im}^* = X_{im}\beta + GE_{im}\beta_{GE} + u_{im} \quad (16a)$$

$$GE_{im} = X_{im}\pi_X + Z_{im}\pi_Z + v_{im} \quad (16b)$$

where π_X and π_Z denote vectors of coefficients in the reduced-form estimation of GE . By assumption, $(u_{im}, v_{im}) \sim N(0, \Sigma)$, Σ denoting the variance-covariance matrix. Similar to the probit model, the IV probit model (16) is estimated by maximum likelihood.

5. Data and descriptive statistics

The main data source used for the empirical analysis is the Brazilian Demographic Census for the year 2000. All worker-level information is derived from these data, and several of the municipal-level variables are aggregated from them as well. The publicly available sample of the Census data includes almost 20 million individual observations, which makes it representative at the level of the municipality. In total, there were 5507 municipalities in the year 2000, with an average population of about 30000 people. The Census data provide detailed information on employment status, earned income and a range of socioeconomic variables. To get the relevant information about municipalities, two data sources are used in addition to aggregates derived from the Census: *Base de Informações Municipais* between 2000 and 2006 and *Perfil dos Municípios Brasileiros - Gestão Pública, 2005–2006*, both from the Brazilian Institute of Geography and Statistics (IBGE). These databases contain detailed information on the structure of the local economy, various demographic characteristics and indicators of public sector management at the municipal level.

5.1. The informal sector in Brazil

The definition of the informal sector used in this paper is based on the employment categories defined in the Demographic Census. Informal employment is defined as being an unregistered

1
2
3 employee, a self-employed person (not contributing to a social security institution), or an employer
4 who has fewer than five employees and does not contribute to any social security institution.⁴ Unpaid
5 workers (who constitute about 5% of the informal sector) are excluded from the empirical analysis.
6
7 Only the urban labour force is analyzed in this study, mainly because the majority of rural residents
8 are engaged in agriculture and do not have access to either a formal or an informal labour market to
9 the same extent as in urban areas. Using this definition, 45% of the urban labour force in Brazil is
10 informal. Approximately 60% are employees and about 40% are self-employed in the informal sector.
11
12

13
14
15
16
17
18 The informal sector in Brazil has some of the characteristics that are commonly observed in
19 studies concerned with informal employment in Latin America and elsewhere. Table 1 provides some
20 key indicators. First, labour incomes are on average considerably lower in the informal than in the
21 formal sector. Average earnings per month in the informal sector are just above 400 Reais per month
22 (about 200 US\$ in year 2000), compared to 769 Reais (R\$) in the formal sector. At the 20th percentile
23 of the earnings distribution in the informal sector, earnings are 120 R\$ – well below the minimum
24 wage (151 R\$) – compared to 221 R\$ at the same percentile in the formal sector. The fact that the
25 earnings at the 20th percentile in the formal sector are reached in the informal sector at approximately
26 the 50th percentile shows that there is a considerable overlap in the earnings distributions in the two
27 sectors. Second, education is lower in the informal sector. On average, a worker in the informal sector
28 has 2.5 years less education than a worker in the formal sector. Similar to the distributions of income,
29 there is an overlap in terms of education between the two sectors. For example, 27% of the workers in
30 the informal sector have nine or more years of education, while almost 50% in the formal sector have
31 less than nine years of education. Third, the share of workers who have just entered the labour market
32 or are close to retirement is higher in the informal sector, even though this is not as pronounced as in
33 many other countries (Perry, et al., 2007). One aspect in which Brazil deviates from many other
34 examples is the gender distribution across the sectors. While the informal sector in many countries
35 tends to be over-represented by women, there is no such gender bias in Brazil. In terms of industrial
36 composition, there is a slight bias towards manufacturing and domestic services in the informal sector
37 compared with the composition of the formal sector.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

There is a vast variation of informality in local labour markets. The average share of informality at municipal level is 55%, but around this mean, informality varies between 20% and 80%. There are more than 300 municipalities with shares below 30% and over 1100 with shares above 70%. This degree of variation in informality is not altered much when small municipalities of 10000 or fewer inhabitants are excluded. Even among the 200 most populated municipalities – each with 400000 or more inhabitants – informality varies between 25% and 70%.

[Table 1]

5.2. Explanatory variables

Table 2 gives an overview of the explanatory variables used in the empirical analysis. Human capital is represented by the individual's age and years of education. Squared age and education are included to test for non-linearity. Two additional variables are added to control for labour productivity: an indicator for physical or mental disability, which is likely to affect work ability, and an indicator showing if the worker has recently migrated from a rural area. It could be that people who have recently entered the urban labour force from rural areas face a disadvantage in terms of knowledge of the local economy and have less access to social networks and informal institutions, which affect labour productivity.

Variables included under the assumption that they affect the valuation of non-pecuniary benefits of informal-sector employment are: gender, position in the household, marital status, and indicators for the presence of formal-sector workers and young children in the household. Gender bias in household-related work and responsibilities may imply that women benefit more from the work flexibility in the informal sector than men. This effect may be strengthened if there are children in the household.

The variable for regional government effectiveness is developed by Jonasson (2011) and is based on the municipal data described above. It is constructed as an index and has similarities with the Brazilian *IQIM* index of local institutional quality used by the Brazilian Ministry of Planning. It also has similarities with the *Worldwide Governance Indicators* (Kaufmann, et al., 2009). A common

1
2
3 feature of these measures is that they are constructed based on a range of indicators in order to obtain
4
5 an index that can rank regions or countries according to quality of governance and institutions.
6

7
8 The composite variable *government effectiveness* (GE) is given by the average of three indices
9
10 that are constructed to capture different aspects of local governance and bureaucracy. The first
11
12 component is *policy formulation* (GE_1), which is a proxy for the capacity of the local administration to
13
14 formulate and implement policy. It is based on 20 indicators, showing to what extent the municipality
15
16 has councils, development plans, ordinances, and regulations in various domains such as education,
17
18 urban development, employment, and property ownership. The second component represents
19
20 *bureaucratic resources* (GE_2) and is based on information about the employment form and
21
22 competence of bureaucrats; the degree of co-ordination between units with different areas of
23
24 responsibility; and information technology resources. The third component represents the quality of,
25
26 and access to, *public goods and services* (GE_3). This index is based on the teacher/pupil ratio in public
27
28 primary schools, the number of health centres per municipality inhabitant, the degree of internet
29
30 services offered to the public, the existence of public libraries, and the degree of support for helping
31
32 people with housing. All index variables are between zero and one. The correlation coefficients of
33
34 GE_1 , GE_2 , and GE_3 , range between 0.24 and 0.42.⁵ The effect of enforcement of tax and labour
35
36 regulation, as discussed in the theoretical model, is not tested for explicitly since it is difficult to
37
38 separate this from other qualities of the local bureaucracy (GE_2).
39
40
41

42
43 A number of municipal control variables are included as well. The sectoral composition of the
44
45 municipal economy is controlled for using the shares of agriculture, manufacturing, services, and
46
47 public sector production in the total municipal gross product. While factors that affect informality
48
49 may influence the structure of the local economy in the long run, the concern here is that the relative
50
51 demand for informal labour may be higher in local economies where certain sectors dominate (such as
52
53 agriculture or services). Average firm size (number of employees) in the formal sector is included to
54
55 control for technology in the formal sector, under the assumption that labour is more productive in
56
57 large firms, which would increase the expected income in the formal sector for any level of worker-
58
59 skill. A distance-weighted measure of local population size is included to control for market potential,
60
assuming that the higher the market potential the higher the returns to acting formally. The share of

1
2
3 rural-to-urban migrants in the labour force is added under the assumption that it will increase the
4 relative supply of unskilled labour and drive down the labour income in the informal sector. Lastly,
5
6
7 municipal product per capita is included to check to what extent the results remain robust while
8
9
10 keeping income level constant.

11 [Table 2]

12 13 14 15 16 5.3 Instrumental variables

17
18 The instrumental-variable approach outlined in Section 4 is used in order to take into account
19 this potential endogeneity of government effectiveness. Instrumental variable candidates are needed
20 that are correlated with *GE*, yet uncorrelated with the error terms of the two models. Naritomi et al.
21 (2007), who use governance as an indicator of institutional development in regions of Brazil, find that
22 a set of geographical variables are significantly related to both economic and institutional
23 development. Almeida and Carneiro (2009), who analyze the effect of labour regulation enforcement
24 on informal employment, suspect that enforcement may be endogenous and use distance to the nearest
25 enforcement office interacted with the local intensity of labour inspectors as an instrumental variable.
26 Based on these studies, three geographical variables were chosen as instruments: longitude, latitude,
27 and transportation cost to state capital from the seat of the municipality. The age of the municipality is
28 included as a fourth instrumental variable. More than 1400 of Brazil's municipalities were created
29 after the constitutional reform in 1988. There is some anecdotal evidence that some of these
30 municipalities were partly created out of rent-seeking motives and that governance performance has
31 developed quite poorly in some of these new municipalities (Economist, 2008; Instituto Brasileiro de
32 Geografia e Estatística, 2001). When the government effectiveness variable, *GE*, is regressed on the
33 instruments and the other exogenous municipal variables, all the instruments show a statistically
34 significant conditional correlation with *GE*.
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

54
55 The validity of the instruments was checked with F tests of the hypothesis that all instrumental
56 variables in an OLS estimation of model (16b) have zero coefficients. The critical F statistic given by
57 Stock et al. (2002) is 9.08 for three instruments and 10.83 for five instruments. The F-statistic for the
58 null hypothesis that the coefficients of the three geographical variables are jointly zero was 16.74. As
59
60

1
2
3 all four variables were tested for joint insignificance, the F-statistic was 22.80. While the instruments
4 suffer from correlation with *GE*, these tests suggest that the instruments bear some validity.
5
6
7
8
9

10 11 12 **6. Empirical results** 13

14 Table 3 contains the results of both the basic probit model and the IV probit model. The binary
15 variable indicating whether the worker is employed in the informal sector (*is*) serves as the dependent
16 variable in all specifications. Six specifications are reported, in which variables have been added
17 stepwise to evaluate the validity of the hypotheses and assess the stability of the coefficient estimates
18 as additional factors are controlled for. The coefficients show marginal effects of small changes in the
19 independent variables or changes from zero to one for binary variables. Standard errors are adjusted to
20 take into account clustering of the error term within municipalities. Due to the large sample size (more
21 than 2 million observations), standard errors are still small after this adjustment and most of the
22 coefficients are statistically significant at very high levels. Specific attention should therefore be given
23 to the economic significance when interpreting the coefficient estimates.
24
25
26
27
28
29
30
31
32
33
34

35 The first column of Table 3 show results from a specification with only worker characteristics
36 included. In line with the model prediction, human capital has a negative effect on the probability of
37 working in the informal sector. Age has a negative but decreasing effect, while years of education
38 appears to have a negative and increasing effect, as indicated by the coefficient estimates of the square
39 terms of these variables. An increase in education by 2.5 years (which is the difference in average
40 education between the two sectors) decreases the probability of informal employment by about six
41 percentage points. This relationship remains stable across the specifications, as municipal-level
42 variables are added.
43
44
45
46
47
48
49
50
51
52

53 In column 2, government effectiveness (*GE*) is included in the model, but without any
54 additional municipal control variables. The coefficient estimate is negative (in line with the
55 predictions of the theoretical model) and statistically significant. The magnitude of -0.24 of the
56 coefficient estimate implies that an increase of the index value of *GE* by one standard deviation, from
57 the average of 0.35 to 0.49, while holding everything else constant, is associated with a decline in the
58
59
60

1
2
3 probability of a worker being informally employed of about 3.5 percentage points. As a series of
4
5 municipal control variables are added to the model, the *GE* coefficient estimate shrinks to about -0.15,
6
7 but still significant at the 1-percent level (column 3).
8

9
10 To gain further insight into how government effectiveness might affect informality, the *GE*
11
12 index is decomposed into its three sub-components and included in the probit model. The results in
13
14 column 4 suggest that resources for policy formulation (GE_1) are less important than the quality of
15
16 bureaucracy (GE_2) and public service provision (GE_3). One interpretation of these results relates to the
17
18 patterns of public trust in politicians, as discussed in the beginning of the paper; political promises
19
20 delivered by means of planning, regulation, and the creation of municipal ordinances may have little
21
22 effect on citizens' incentives if they doubt that these political efforts will have any real effect on them
23
24 (Saavedra and Tommasi, 2007). Quality of the bureaucracy and public services, on the other hand,
25
26 might have a more direct effect on incentives in terms of actual enforcement of regulation and other
27
28 value-added in the formal sector.
29
30

31
32 The only case in which the government effectiveness coefficient becomes statistically
33
34 indistinguishable from zero is when municipal gross product per capita is included as a control
35
36 variable (column 5). This suggests that the level of economic development has a more important
37
38 effect on informality than government effectiveness. However, income per capita should probably be
39
40 perceived, to a larger extent than any other control variable, as endogenously determined by other
41
42 municipal characteristics. This potential endogeneity of the local income level calls for caution in the
43
44 interpretation of the results when local income is included.⁶
45

46
47 Column 6 reports the results from one of all the IV probit specifications that were tested. This
48
49 specification is based on the specification in column 3, with the difference that government
50
51 effectiveness is instrumented for using the four instrumental variables discussed above. Exogeneity of
52
53 *GE* in this IV probit specification was rejected by a Wald test at the one-percent level, confirming the
54
55 suspicion of endogeneity. The *GE* coefficient remains negative and statistically significant, even
56
57 though it is larger in magnitude (-0.46 instead of -0.15 in column 3). Experiments with a series of
58
59 other combinations of instrumental variables gave qualitatively similar results, even though the
60
magnitude of the coefficients varied with specification.

[Table 3]

A few observations might be warranted regarding the coefficients in Table 3 that are not part of the core hypotheses. Ten individual characteristics other than human capital proxies are included in the analysis. First, being female has a very small positive effect on the probability of being in the informal sector. This gender effect is notable only where there are young children in the household (according to the interaction coefficient). Among the coefficients of the local economy control variables (z), it can be noted that the share of manufacturing in the local economy has a negative effect on informality. Except in one case, the share of agriculture and the relative size of the service sector show no significant relationship with informality (the excluded sector being the public sector). There is a positive relationship between informality and the share of the urban labour force that consists of recent rural-to-urban migrants (even when controlling for migrant status of the individual). Rural migrants are likely to increase the relative supply of low-skilled labour, drawn primarily to the informal sector. Lastly, the size of the local population has a small negative effect on informality, giving some support to the notion that there are higher returns to formality in larger markets. This could be due to greater opportunities for specialisation as well as economies of scale in production. Robustness checks discussed in the Online Appendix show that the coefficient estimates of main interest generally remain stable and significant with the expected sign. In cases where the sample size is reduced to a small subset, some deviation in the results is observed.

7. Conclusion

The aim of this paper is to assess how informal employment is affected by the extent to which the government supports and implements formal institutions and provides high-quality public goods and services, referred to here as government effectiveness. A simple theoretical model was outlined to discuss how this may be directly related to worker decisions and informal employment. The empirical assessment of the model supports the main hypotheses: informality is higher where education is lower

1
2
3 and where governance is less effective. These results complement previous studies by showing that
4 regional factors, which are exogenous to the individual, affect individual employment outcomes. They
5 also complement cross-country studies by showing that regional differences at sub-national level may
6
7
8
9
10
11
12
13 cause informality to vary just as much as between countries, despite the fact that the formal
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60 institutional framework is held constant.

The ambition of the paper is not to provide specific policy advice on which actions to take in order to reduce informal employment. Nevertheless, a conclusion from the empirical results is that human capital investment is the principal means of increasing the chances of workers to find formal employment and strengthening the incentives to choose formal employment. But the results also suggest that the incentive structure may go beyond economic returns to formalisation. For the local government, the most challenging task – besides providing education, efficient bureaucracy and other public services to its citizens – might be to improve the quality of the “social contract” between the authorities and citizens. This implies making participation in the formal sector the norm rather than an exception. While such norms are likely to change only slowly over time, a government can seek ways to improve the sense of political participation and inclusion among its citizens. This includes transparency in the political decision process and in the spending of public resources, as well as recognizing the needs of the people outside the formal sector just as much as the needs of those who are already in it.

Endnotes

¹ The assumption that the probability of finding a job in the formal sector depends on worker skill stems back to Fields (1975), who assumes that workers are employed in the “murky” (informal) sector while searching for jobs in the urban (formal) sector.

² Self-employed workers do not receive labour income from an employer, but get paid for their production of goods and services. For simplicity, assume that production among self-employed is given by $q = L^{\beta}(s)$ and that the price of output equals one.

³ Both in Figure 1 and Figure 2 the parameters α and β are assumed, for simplicity, to sum to 1; hence the linearity of y^{FS} .

⁴ Henley et al. (2009) provide an analysis of three alternative measures of informal employment in Brazil, using the PNAD survey. The definition used here corresponds to a large extent to their measure of informality defined as “no signed labor card”.

⁵ As noted by Langbein and Knack (2010), the six *Worldwide Governance Indicators* show bivariate correlations between 0.64 and 0.91.

⁶ This is discussed as part of several robustness checks in the Online Appendix.

References

- Almeida, R. and Carneiro, P. (2009) Enforcement of Labor Regulation and Firm Size. *Journal of Comparative Economics*, 37(1), pp. 28-46.
- Bosch, M., Goni, E., and Maloney, W. (2007) The Determinants of Rising Informality in Brazil: Evidence from Gross Worker Flows. Washington, DC: *World Bank Policy Research Working Paper 4375*. The World Bank.
- Casson, M., Della Giusta, M., and Kambhampati, U. (2010) Formal and Informal Institutions and Development. *World Development*, 38(2), pp. 137-141.
- Chaudhuri, K., Schneider, F., and Chattopadhyay, S. (2006) The Size and Development of the Shadow Economy: An Empirical Investigation from States of India. *Journal of Development Economics*, 80(2), pp. 428-443.
- de Paula, Á. and Scheinkman, J. (2010) Value-Added Taxes, Chain Effects, and Informality. *American Economic Journal: Macroeconomics*, 2 (4), pp. 195-221.
- Economist. (2008) Real-politik: Why Elections, Even to Lowly Office, Are so Expensive. *Economist*, 388, September 27, pp 50-52.
- Fajnzylber, P., Maloney, W., and Montes-Rojas, G. V. (2009) Releasing Constraints to Growth or Pushing on a String? Policies and Performance of Mexican Micro-Firms. *Journal of Development Studies*, 45(7), pp. 1027-1048.
- Fields, G. (1975) Rural-Urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs. *Journal of Development Economics*, 2(2), pp. 165-187.
- Freeman, R. (2009) Labor Regulations, Unions, and Social Protection in Developing Countries: Market Distortions or Efficient Institutions? Cambridge, MA: *NBER Working Paper Series 14789*. National Bureau of Economic Research.
- Friedman, E., Johnson, S., Kaufmann, D., and Zoido-Lobaton, P. (2000) Dodging the Grabbing Hand: the Determinants of Unofficial Activity in 69 Countries. *Journal of Public Economics*, 76(3), pp. 459-493.
- Funkhouser, E. (1996) The Urban Informal Sector in Latin America: Household Survey Evidence. *World Development*, 24(11), pp. 1737-1751.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Galiani, S. and Weinschelbaum, F. (2010) Modeling Informality Formally: Households and Firms. La Plata: CEDLAS Working paper. Centro de Estudios Distributivos, Laborales y Sociales.
- Glaeser, E., La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (2004) Do Institutions Cause Growth? *Journal of Economic Growth*, 9(3), pp. 271-303.
- Gong, X., van Soest, A., and Villagomez, E. (2004) Mobility in the Urban Labor Market: A Panel Data Analysis for Mexico. *Economic Development and Cultural Change*, 53(1), pp. 1-36.
- Hall, R. and Jones, C. (1999) Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, 114(1), pp. 83-116.
- Hansen, H., Rand, J., and Tarp, F. (2009) Enterprise Growth and Survival in Vietnam: Does Government Support Matter? *Journal of Development Studies*, 45(7), 1048-1069.
- Henley, A., Arabsheibani, R., and Carneiro, F. (2009) On Defining and Measuring the Informal Sector: Evidence from Brazil. *World Development*, 37(5), pp. 992-1003.
- Instituto Brasileiro de Geografia e Estatística. (2001, April 17). IBGE traça o perfil administrativo de todos os municípios brasileiros. Retrieved June 2, 2011, from <http://www1.ibge.gov.br/home/presidencia/noticias/1704munic.shtm>
- Johnson, S., Kaufmann, D., and Zoido-Lobaton, P. (1998) Regulatory Discretion and the Unofficial Economy. *American Economic Review*, 88(2), pp. 387-392.
- Jonasson, E. (2011) Informal Employment and the Role of Regional Governance. *Review of Development Economics*, forthcoming.
- Jütting, J., Parlevliet, J., and Xenogiani, T. (2008) Informal Employment Re-loaded. Paris: *OECD Development Centre Working Paper 266*. OECD Development Centre.
- Kaufmann, D., Kraay, A., and Mastruzzi, M. (2009) Governance Matters VIII: Aggregate and Individual Governance Indicators 1996–2008. Washington, DC: *Policy Research Working Paper 4978*. World Bank.
- Kucera, D. and Xenogiani, T. (2009) Persisting Informal Employment: What Explains It? In J. Jütting & J. de Laiglesia (eds), *Is Informal Normal? Towards More and Better Jobs in Developing Countries*, pp. 63-88. (Paris: OECD)
- Langbein, L. and Knack, S. (2010) The Worldwide Governance Indicators: Six, One, or None? *Journal of Development Studies*, 46(2), pp. 350-370.
- Loayza, N. (1996) The Economics of the Informal Sector: A Simple Model and Some Empirical Evidence from Latin America. *Carnegie-Rochester Conference Series on Public Policy*, 45, pp. 129-162.
- Loayza, N., Oviedo, A. M., and Servén, L. (2006) The Impact of Regulation on Growth and Informality: Cross-Country Evidence. In B. Guha-Khasnobis, R. Kanbur, and E. Ostrom (eds), *Linking the Formal and Informal Economy: Concepts and Policies*, pp. 121-144. (Oxford: Oxford University Press)

- 1
2
3 Loayza, N. and Rigolini, J. (2011) Informal Employment: Safety Net or Growth Engine? *World*
4 *Development*, forthcoming.
5
6 Loayza, N., Servén, L., and Sugawara, N. (2010) Informality in Latin America and the Caribbean. In
7 N. Loayza and L. Servén (eds), *Business Regulation and Economic Performance: A Latin*
8 *American Perspective*, pp. 157-196. (Washington, DC: The World Bank)
9
10 Maloney, W. (2004) Informality Revisited. *World Development*, 32(7), pp. 1159-1178.
11
12 Marcouiller, D., Ruiz de Castilla, V., and Woodruff, C. (1997) Formal Measures of the Informal-
13 Sector Wage Gap in Mexico, El Salvador, and Peru. *Economic Development and Cultural*
14 *Change*, 45(2), pp. 367-392.
15
16 Naritomi, J., Soares, R., and Assunção, J. (2007) Rent Seeking and the Unveiling of 'de facto'
17 Institutions: Development and Colonial Heritage within Brazil. Cambridge, MA: *NBER Working*
18 *Paper 13545*. National Bureau of Economic Research.
19
20 Perry, G., Maloney, W., Arias, O., Fajnzylber, P., Mason, A., and Saavedra-Chanduvi, J. (2007)
21 *Informality: Exit and Exclusion*. Washinton, D.C.: The World Bank.
22
23 Rauch, J. (1991) Modelling the Informal Sector Formally. *Journal of Development Economics*, 35 (1),
24 pp. 33-47.
25
26 Saavedra, J. and Tommasi, M. (2007) Informality, the State and the Social Contract in Latin America:
27 A Preliminary Exploration. *International Labour Review*, 146(3-4), pp. 279-309.
28
29 Schneider, F. and Enste, D. (2000) Shadow Economies: Size, Causes, and Consequences. *Journal of*
30 *Economic Literature*, 38(1), pp. 77-114.
31
32 Stock, J., Wright, J., and Yogo, M. (2002) A Survey of Weak Instruments and Weak Identification in
33 Generalized Method of Moments. *Journal of Business and Economic Statistics*, 20(4), pp. 518-
34 529.
35
36 Straub, S. (2005) Informal Sector: The Credit Market Channel. *Journal of Development Economics*,
37 78(2), pp. 299-321.
38
39 Torgler, B. and Schneider, F. (2007) Shadow Economy, Tax Morale, Governance and Institutional
40 Quality: A Panel Analysis. Bonn: *IZA Discussion Paper Series 2563*. Institute for the Study of
41 Labor.
42
43 Torgler, B. and Schneider, F. (2009) The impact of tax morale and institutional quality on the shadow
44 economy. *Journal of Economic Psychology*, 30(2), pp. 228-245.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. The urban labour market in Brazil

		<i>Age (percent of labour force in each age category)</i>					
		15-24	25-34	35-44	45-54	55-65	average age
Formal sector		21	31	27	16	5	35
Informal sector		28	27	23	15	7	34
		<i>Years of education (percent of labour force in each category)</i>					
		0-2	3-5	6-8	9-12	>12	average years
Formal sector		8	21	19	36	16	8.6
Informal sector		19	32	22	22	5	6.1
		<i>Earnings, at percentile (R\$ per month)</i>					
		20%	40%	60%	80%		average
Formal sector		221	330	500	970		769
Informal sector		120	167	300	500		404
		<i>Sector of employment (percent per sector)</i>					
		Commerce	Manufacturing	Construction	Domestic services		Other services
Formal sector		21	13	12	5		46
Informal sector		18	18	5	14		30

Note: In August 2000, the exchange rate was R\$1 = US\$0.56. Source: Brazilian Demographic Census, 2000.

Table 2. Variables used in the empirical analysis

Variable	Mean	S.D.	Definition
Worker level (N = 2,222,387)			
Informal employment	0.436	0.496	Individual works in informal sector, with earned non-zero income
Age	34.5	11.6	Individual's years of age
Education	7.5	4.3	Individual's years of schooling
Gender - female	0.397	0.489	Individual is female
Household head	0.516	0.500	Individual is household head
Married	0.454	0.498	Individual is married
Race - black	0.066	0.248	Individual is black
Disabled	0.021	0.144	Individual has reported physical or mental problem
Rural-urban migrant	0.029	0.169	Individual has moved recently from a rural area
Children in household	0.480	0.500	There is at least one child in the household, <10 years old
Municipal level (N = 5,506)			
Share informal employment	0.554	0.161	Share of the urban municipal labour force that has informal employment.
Average education	6.4	1.1	Average years of education in the municipal labour force.
<i>Municipal variables</i>			
Government effectiveness, GE	0.348	0.139	Index composed of GE_1 , GE_2 , and GE_3 , below, to represent government effectiveness in the municipality.
Policy formulation, GE_1	0.253	0.183	Index composed of 20 variables to represent the capacity of the municipal government to formulate and implement policy.
Bureaucratic resources, GE_2	0.427	0.131	Index that represents the human, technical, and managerial resources available to the municipal bureaucracy.
Public goods, GE_3	0.363	0.234	Index composed of 10 variables to represent the quality of, and access to, public goods in the municipality.
Share agriculture	0.232	0.163	Agriculture as share of gross municipal product.
Share manufacturing	0.163	0.157	Manufacturing as share of gross municipal product.
Share services	0.399	0.094	Services as share of gross municipal product.
Share public sector	0.207	0.084	Public sector production share of gross municipal product.
Average firm size	5.4	5.7	Average number of employees in registered businesses.
Local population	694	1,464	Local population estimate (thousands), weighted by distance
Gross product per capita	4,435	5,699	Municipal gross product per capita, R\$, year 2000.
Share rural immigrants	0.056	0.047	Share of the urban labour force consisting of workers who have migrated from a rural area, five years or less prior to the survey.
<i>Instrumental variables</i>			
Latitude	-16	8	Latitude of municipal seat
Longitude	-46	6	Longitude of municipal seat
Transport cost to state capital	469	409	Transportation cost from municipal seat to state capital city
Age of municipality	55	57	Years since the municipality was created

Sources: Individual-level variables and informal employment and labour force skill on municipal level – Brazilian Demographic Census, 2000. Municipal variables for government effectiveness and the local economy – Gestão Pública, 2005–2006, and Base de Informações Municipais, 2000.

Table 3. Empirical results: probability of informal employment.

	(1)	(2)	(3)	(4)	(5)	(6) IV
<i>Human capital (s)</i>						
Age	-0.023***	-0.022***	-0.022***	-0.022***	-0.022***	-0.022***
Age, squared	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
Education	-0.029***	-0.026***	-0.024***	-0.024***	-0.022***	-0.022***
Education, squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.001***	-0.000***
<i>Other individual characteristics (x)</i>						
Gender - female	0.007**	0.008***	0.008***	0.008***	0.007**	0.009***
Household head	-0.031***	-0.031***	-0.033***	-0.033***	-0.033***	-0.032***
Married	-0.052***	-0.056***	-0.053***	-0.052***	-0.051***	-0.053***
Race - black	-0.012***	-0.008***	-0.007**	-0.006**	-0.008**	-0.007**
Disabled	0.052***	0.050***	0.048***	0.048***	0.046***	0.045***
Rural-urban migrant	0.030***	0.020***	0.009**	0.009**	0.013***	0.013**
Children in household	0.018***	0.015***	0.012***	0.012***	0.008***	0.009***
Female × child in hhd	0.047***	0.047***	0.046***	0.046***	0.047***	0.046***
<i>Government effectiveness</i>						
Government effectiveness, GE		-0.244***	-0.153***		0.004	-0.463**
Policy formulation, GE_1				-0.021		
Bureaucratic resources, GE_2				-0.114***		
Public goods, GE_3				-0.047***		
<i>Municipal control variables (z)</i>						
Gross product per capita					-0.091***	
Share agriculture			-0.116	-0.106	-0.041	-0.368*
Share manufacturing			-0.283***	-0.284***	-0.165**	-0.399***
Share services			-0.022	-0.019	-0.120**	-0.136
Average firm size			-0.000	-0.000	0.002	0.002
Share rural immigrants			0.263*	0.253*	0.247**	-0.067
Local population			-0.018***	-0.018***	-0.010***	-0.016***
Sample size	2,220,387	2,218,167	2,213,429	2,213,429	2,213,429	2,213,429
McFadden pseudo R^2	0.0728	0.0785	0.0846	0.0848	0.0890	

Note: Dependent variable is the dummy variable indicating informal employment (is). Coefficients show marginal effects. Asterisks denote level of significance: *** 1%, ** 5%, and * 10%. Standard errors are adjusted for error term clustering. Coefficient estimates that are reported as "0.000", and yet statistically significant, are smaller than 0.0005 in absolute magnitude. Column 6 contains results from an IV specification, in which GE is instrumented for.

Figure 1. Skill thresholds equating expected earnings in the formal and informal sectors

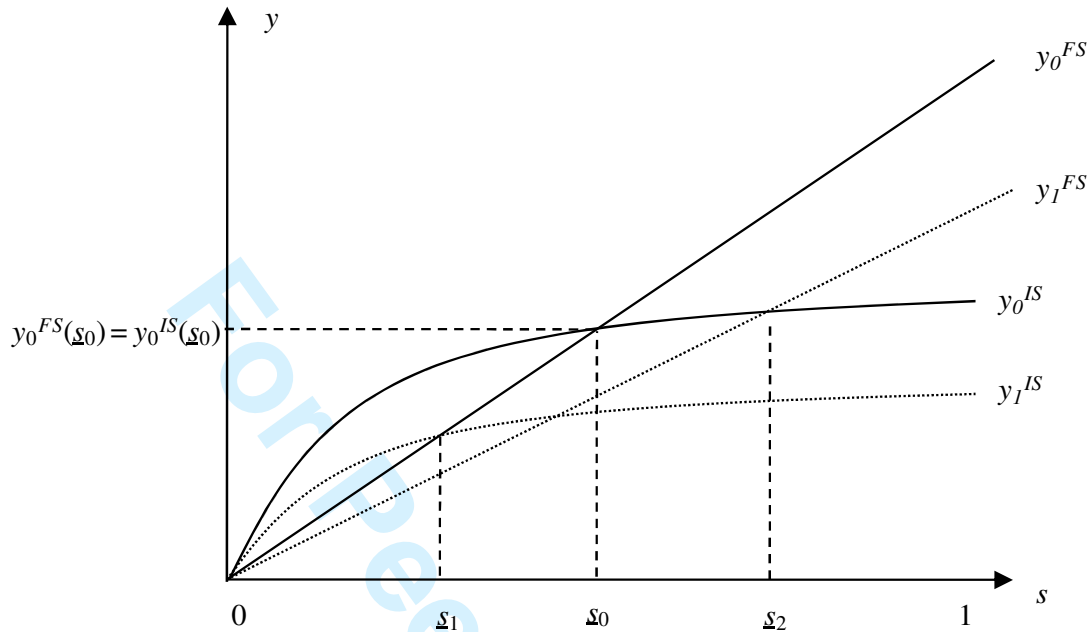
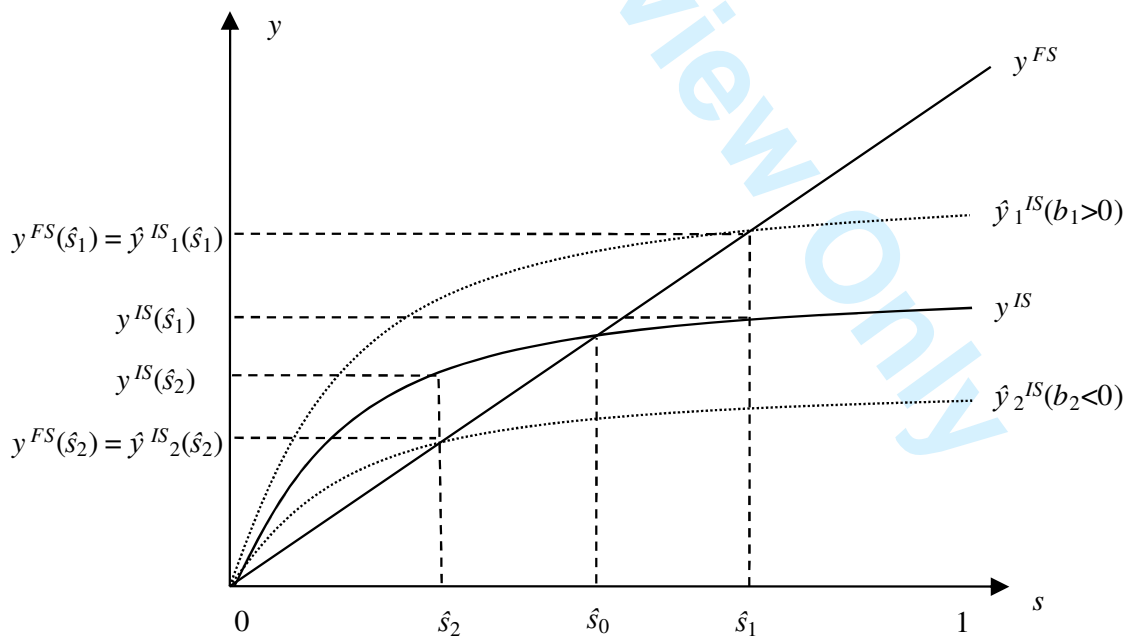


Figure 2. Skill thresholds in the presence of non-pecuniary costs and benefits in the informal sector



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

**Online Appendix to
“Government Effectiveness and Regional Variation in Informal Employment”:**

Robustness of the empirical results

Some further endogeneity concerns and limitations of the empirical results are elaborated on briefly in this appendix. First, per-capita income, as a general indicator of local economic development, is likely to affect – and be affected by – several observable and unobservable variables (including quality of governance). Due to this potential endogeneity, it is included only in one specification of each model. This resulted in a statistically insignificant coefficient estimate of the governance index. To assess the importance of the municipal key variables while still holding local per-capita income “constant”, the models were evaluated on a sub-sample of municipalities with relatively homogenous income. A 40-percent sub-sample of municipalities, consisting of the approximately 2 100 “middle-income municipalities” with per-capita income of between 2 000 and 5 000 R\$, was selected for this purpose. While coefficient estimates change slightly in magnitude, no qualitative changes occur with this sub-sample. A tentative conclusion is that the results are not driven to any large extent by differences in productivity or per-capita income.

Second, the structure of the local economy could be endogenously determined in the model, just as governance might be. The structure of the local economy could be affected by human capital intensity, institutional and economic development, geography, or by unobserved characteristics. The models were therefore evaluated on sub-samples with relatively homogenous structures of the local economy. While municipalities dominated by services and manufacturing do not deviate from the previously obtained results, agriculture-dominated municipalities do not show the same strong relationships between informality and governance.

Third, the empirical literature on spatial human capital externalities is usually concerned about endogenous sorting of skilled people to certain regions (Moretti, 2004). Thus the average level of human capital in a city or region might not be exogenously determined. While beyond the scope of this study to

1
2
3 fully adjust for this possible endogeneity (by the means of additional instrumental variables), the method
4 of evaluating the models on sub-samples has been applied. By using the sub-sample of individuals who
5 have never moved from one municipality to another, some of the endogenous-sorting problem is taken
6 care of. Even if the resulting sample size decreases by half, the key coefficient estimates remain robust,
7 with only minor changes in magnitude. Another sub-sample includes only municipalities with relatively
8 homogenous education (those less than half a standard deviation away from the average level of
9 education). Parameter estimates from regressions on this sample are similar to those on the full sample,
10 with the exception for two of the disaggregated governance indices in some of the specifications.
11
12
13
14
15
16
17
18
19

20
21 Fourth, in the results discussed above there is no distinction made between the self-employed and
22 employees in the informal sector. While these two groups are treated as homogenous in this study, they
23 might face different income prospects and hence different incentives regarding sector choice. Some
24 empirical studies focus solely on employees (Pratap and Quintin, 2006) or self-employed (Blau, 1985) in
25 the informal sector, due to their potentially fundamental differences. No major deviations in the results
26 are observed when the probit model is estimated on sub-samples with a) all informal employees excluded
27 and b) with all informal self-employed excluded.
28
29
30
31
32
33
34
35

36 In sum, the robustness checks show that the coefficient estimates of main interest generally remain
37 stable and significant with the expected sign. In cases where the sample size is reduced to a small subset,
38 some deviation in the results is observed.
39
40
41
42
43
44

45 **References**

- 46 Blau, D. (1985) Self-Employment and Self-Selection in Developing Country Labor Markets. *Southern*
47 *Economic Journal*, 52(2), pp. 351-363.
48
49 Moretti, E. (2004) Human Capital Externalities in Cities. In V. Henderson and J.-F. Thisse (eds),
50 *Handbook of Regional and Urban Economics*, Vol. 4, pp. 2243-2291. (Amsterdam: Elsevier
51 Science)
52
53
54 Pratap, S. and Quintin, E. (2006) Are labor markets segmented in developing countries? A
55 semiparametric approach. *European Economic Review*, 50(7), pp. 1817-1841.
56
57
58
59
60