**ONLINE APPENDICES**

**Appendix A – Reweighting and trimming**

In line with best practice, we estimate the propensity of each household to be linked over 1, 5 and 10 year transition periods using a series of logit regressions. The results are shown in Table A1. Panel A shows a baseline specification, which only includes annual fixed effects, and an indicator for whether the taxed household head has a wife present in the household. This coefficient is allowed to vary over period and district. As expected, a present wife increases matching probabilities significantly, though the relationship grows weaker over time. The squared correlation between predicted propensity scores and the linkage indicator is 0.235 for Stellenbosch households in a 1 year transition window, and somewhat stronger at 0.271 for Graaff-Reinet. These fit statistics are somewhat weaker for predicting matches at 5 and 10 year horizons. Panel B extends this specification by adding a number of control variables for which the outputs are omitted, viz. *log(children in household), log(horses), log(cattle), log(name frequency), log(name frequency)2, log(vines), log(enslaved)*. Each of these variables is also interacted with period indicators. Notably, whether a wife is present remains a robust predictor of linkage, and the fit statistics do not improve substantially when we add the control variables. Therefore, the best predictor of finding a link is whether two names were available in the records for matching – that of a male and a female. While we do not have an appropriate validation variable in our data, we use the full logit specifications to generate inverse propensity score weights for reweighting and trimming. We generated weights that were specific to each of the transition windows. We trimmed observations with propensity scores below 0.3, the average for 5 year transitions. Our results were robust to truncating observations with propensity scores above 0.5 in the case of 1 and 5 year transitions. However, because propensity scores are generally much lower for 10 year transitions, our sample becomes restrictively small if we use a higher threshold in that instance. We therefore base all our analysis on a threshold of 0.3.

**Table A1: Logit specifications of linkage over various transition periods and districts**

|  |  |
| --- | --- |
|  | 1. *Simpler specification*
 |
|  | (1) | (2) | (3) |  | (4) | (5) | (6) |
|  | Stellenbosch |  | Graaff-Reinet |
| Transition | 1 year | 5 year | 10 year |  | 1 year | 5 year | 10 year |
|  |  |  |  |  |  |  |  |
| Wife present  | 0.504\*\*\* | 0.526\*\*\* | 0.447\*\*\* |  | 1.027\*\*\* | 0.923\*\*\* | 1.717\*\*\* |
| (period 1) | (0.032) | (0.032) | (0.047) |  | (0.076) | (0.120) | (0.055) |
|  |  |  |  |  |  |  |  |
| (period 2) | 0.321\*\*\* | 0.334\*\*\* | 0.433\*\*\* |  | 0.589\*\*\* | 1.029\*\*\* | 0.077 |
|  | (0.054) | (0.060) | (0.077) |  | (0.091) | (0.145) | (0.129) |
|  |  |  |  |  |  |  |  |
| (period 3) | 0.298\*\*\* | 0.365\*\*\* | 0.451\*\*\* |  | 0.435\*\*\* | 0.795\*\*\* |  |
|  | (0.041) | (0.041) | (0.057) |  | (0.084) | (0.126) |  |
|  |  |  |  |  |  |  |  |
| Constant | -0.110\* | -0.739\*\*\* | -1.491\*\*\* |  | -0.297\*\*\* | -1.226\*\*\* | -2.247\*\*\* |
|  | (0.048) | (0.048) | (0.055) |  | (0.062) | (0.064) | (0.076) |
| Year FE | Y | Y | Y |  | Y | Y | Y |
| Controls | N | N | N |  | N | N | N |
| BIC | 72782.29 | 65198.31 | 43614.33 |  | 35848.24 | 33186.27 | 23228.24 |
| pseudo-R2 | 0.049 | 0.039 | 0.037 |  | 0.113 | 0.097 | 0.076 |
| $$r\_{link;\hat{p}}^{2}$$ | 0.235 | 0.164 | 0.132 |  | 0.271 | 0.197 | 0.154 |
| N | 55678 | 51231 | 42243 |  | 30149 | 26984 | 21749 |
|  |  |
|  | *B. Fuller specification* |
|  |  |  |  |  |  |  |  |
| Wife present  | 0.504\*\*\* | 0.352\*\*\* | 0.158\* |  | 0.580\*\*\* | 0.355\* | 1.013\*\*\* |
| (period 1) | (0.043) | (0.043) | (0.063) |  | (0.100) | (0.149) | (0.065) |
|  |  |  |  |  |  |  |  |
| (period 2) | 0.160\* | 0.104 | 0.252\* |  | 0.616\*\*\* | 0.823\*\*\* | -0.178 |
|  | (0.073) | (0.080) | (0.101) |  | (0.121) | (0.179) | (0.151) |
|  |  |  |  |  |  |  |  |
| (period 3) | 0.076 | 0.185\*\*\* | 0.354\*\*\* |  | 0.401\*\*\* | 0.654\*\*\* |  |
|  | (0.055) | (0.056) | (0.076) |  | (0.110) | (0.157) |  |
|  |  |  |  |  |  |  |  |
| Constant | -2.121\*\*\* | -3.000\*\*\* | -4.642\*\*\* |  | -1.720\*\*\* | -3.549\*\*\* | -5.233\*\*\* |
|  | (0.140) | (0.181) | (0.268) |  | (0.185) | (0.211) | (0.287) |
| Year FE | Y | Y | Y |  | Y | Y | Y |
| Controls  | Y | Y | Y |  | Y | Y | Y |
| BIC | 72281.203 | 64186.72 | 42344.38 |  | 35300.33 | 31762.52 | 21922.47 |
| pseudo-R2 | 0.059 | 0.057 | 0.071 |  | 0.132 | 0.142 | 0.134 |
| $$r\_{link;\hat{p}}^{2}$$ | 0.245 | 0.184 | 0.163 |  | 0.283 | 0.228 | 0.197 |
| N | 55678 | 51231 | 42243 |  | 30149 | 26984 | 21749 |

**NOTES:** Standard errors in parentheses. Panel A includes only year fixed effects and allows the parameter on *Wife Present* to vary by period. In Stellenbosch period 1 corresponds to 1780-1798, period 2 to 1803-1811 and period 3 to 1812-1829. In Graaff-Reinet period 1 corresponds to 1792-1800, period 2 to 1805-1813 and period 3 to 1814-1823. Panel B includes additional control variables, all of which vary by period: *log(children in household), log(horses), log(cattle), log(name frequency), log(name frequency)2, log(vines), log(enslaved)*.

\* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

 **Figure A1: One year transition plots pooled over time: unweighted, weighted and trimmed**

**NOTES**: Inverse propensity weights were generated from columns (1) and (4) in Table A1. The first column of figures presents unweighted estimates, the second column uses the same sample as in the first column, but reweights, while the third column reweights, but also limits the sample to observations with propensity scores above 0.3.

Figure A1 shows that weighting the full sample has a negligible impact on mobility plots of one year transitions relative to unweighted estimates. Estimates of corresponding elasticities are also barely affected. Trimming the sample to include only observations with propensity scores above 0.3 also has no meaningful effect on conclusions. A similar scenario holds for 5 year transitions, as revealed in Figure A2.

** Figure A2: Five year transition plots pooled over time: unweighted, weighted and trimmed**

**NOTES**: Inverse propensity weights were generated from columns (2) and (5) in Table A1. The first column of figures presents unweighted estimates, the second column uses the same sample as in the first column, but reweights, while the third column reweights, but also limits the sample to observations with propensity scores above 0.3

However, Figure A3 shows that estimates using 10 year matches are more sensitive, as may be expected. Trimming the Stellenbosch sample to propensity scores above 0.3 produces lower estimates of persistence. Without these adjustments, the majority of households in the lowest decile remain in that position 10 years later. However, when the sample is trimmed, significant long-run upward mobility emerges for households initially in the lowest decile. The rest of the distribution shows similar dynamics to the earlier analysis. We therefore trim our sample in our core set of analyses.



**Figure A3: Ten year transition plots pooled over time: unweighted, weighted and trimmed**

**NOTES**: Inverse propensity weights were generated from columns (3) and (6) in Table A1. The first column of figures presents unweighted estimates, the second column uses the same sample as in the first column, but reweights, while the third column reweights, but also limits the sample to observations with propensity scores above 0.3

**Appendix B – Descriptive Statistics**

Table B1: Descriptive statistics

|  |  |
| --- | --- |
| District  | Stellenbosch  |
| Period  | 1780-1798  |
| Statistic  | N  | mean  | min  | p25  | p50  | p75  | max  | % zeroes  |
| Total income (all sources) | 22781  | 381.513  | 1.157  | 2.183  | 63.722  | 338.102  | 16740.7  | 0.05  |
| Crops  | 1722  | 396.31  | 1.003  | 101.765  | 220.815  | 487.93  | 6194.746  | 0.928  |
| Alcohol  | 3424  | 1138.14  | 46.666  | 466.661  | 933.322  | 1399.982  | 9333.216  | 0.857  |
| Animals  | 11656  | 69.286  | 0.415  | 23.132  | 45.072  | 87.761  | 1221.353  | 0.514  |
| Number of enslaved workers  | 10984  | 8.963  | 1  | 3  | 6  | 12  | 198  | 0.542  |
| Period  | 1803-1811  |
| Statistic  | N  | mean  | min  | p25  | p50  | p75  | max  | % zeroes  |
| Total income (all sources) | 9775  | 998.199  | 0.415  | 51.328  | 271.441  | 1043.531  | 301174.3  | 0.253  |
| Crops  | 3449  | 466.126  | 1.752  | 107.722  | 276.512  | 588.761  | 6418.474  | 0.735  |
| Alcohol  | 2375  | 1507.351  | 46.666  | 485.673  | 1045.666  | 1904.667  | 168137.9  | 0.818  |
| Animals  | 6576  | 60.506  | 0.104  | 17.222  | 35.86  | 71.183  | 3173.594  | 0.497  |
| Number of enslaved workers  | 7001  | 9.851  | 1  | 3  | 7  | 13  | 122  | 0.465  |
| Period  | 1812-1829  |
| Statistic  | N  | mean  | min  | p25  | p50  | p75  | max  | % zeroes  |
| Total income (all sources) | 20960  | 957.567  | 0.104  | 30.77  | 219.754  | 1219.429  | 178025.3  | 0.298  |
| Crops  | 6310  | 358.447  | 0.002  | 47.301  | 145.242  | 405.951  | 45531.5  | 0.787  |
| Alcohol  | 7093  | 1605.805  | 46.666  | 485.673  | 1138.998  | 2157.009  | 177471.1  | 0.762  |
| Animals  | 10798  | 47.041  | 0.104  | 16.192  | 32.056  | 61.299  | 730.683  | 0.638  |
| Number of enslaved workers  | 14730  | 9.678  | 1  | 3  | 6  | 13  | 117  | 0.507  |
| District  | Graaff-Reinet  |
| Period  | 1792-1800  |
| Statistic  | N  | mean  | min  | p25  | p50  | p75  | max  | % zeroes  |
| Total income (all sources) | 5211  | 201.033  | 0.104  | 33.268  | 91.05  | 220.024  | 5264.593  | 0.146  |
| Crops  | 570  | 200.535  | 1.343  | 54.531  | 131.046  | 255.664  | 2455.077  | 0.907  |
| Alcohol  | 124  | 154.37  | 2.917  | 46.666  | 93.332  | 186.664  | 1213.318  | 0.98  |
| Animals  | 4429  | 157.148  | 0.104  | 45.452  | 89.136  | 188.18  | 3272.145  | 0.274  |
| Number of Slaves  | 1421  | 3.728  | 1  | 1  | 3  | 5  | 61  | 0.767  |
| Period  | 1805-1813  |
| Statistic  | N  | mean  | min  | p25  | p50  | p75  | max  | % zeroes  |
| Total income (all sources) | 10097  | 232.129  | 1.037  | 42.336  | 111.745  | 263.125  | 9071.587  | 0.063  |
| Crops  | 895  | 318.643  | 3.562  | 140.037  | 246.123  | 401.165  | 5393.339  | 0.917  |
| Alcohol  | 245  | 297.044  | 11.667  | 93.332  | 168.084  | 345.674  | 3615.322  | 0.977  |
| Animals  | 9434  | 153.137  | 0.519  | 42.247  | 98.865  | 201.101  | 1581.457  | 0.124  |
| Number of enslaved workers   | 3229  | 4.059  | 1  | 1  | 2  | 5  | 35  | 0.7  |
| Period  | 1814-1828  |
| Statistic  | N  | mean  | min  | p25  | p50  | p75  | max  | % zeroes  |
| Total income (all sources) | 17854  | 202.221  | 0.519  | 28.086  | 89.061  | 236.287  | 8432.069  | 0.089  |
| Crops  | 2530  | 201.428  | 0.219  | 68.545  | 135.338  | 262.327  | 3326.996  | 0.87  |
| Alcohol  | 1062  | 214.762  | 5.833  | 60.709  | 130.924  | 242.836  | 6756.852  | 0.946  |
| Animals  | 15350  | 126.259  | 0.104  | 32.588  | 78.424  | 167.012  | 1634.267  | 0.217  |
| Number of enslaved workers  | 5156  | 4.376  | 1  | 1  | 3  | 6  | 57  | 0.737  |

NOTES: Own calculations from *Opgaafrollen*. Households are included in all periods they are observed, not only in the times that they are linked across time

**Appendix C – Robustness checks**

****

**Figure C1: Five year transition plots pooled over time by district: including and excluding households with zero income, weighted and trimmed**

**NOTES**: Estimates are weighted and trimmed using inverse propensity weights that were generated from columns (2) and (5) in Table A1. The first column of figures presents estimates that exclude zero incomes, while the second column introduces them into the sample.



**Figure C2: Five year transition plots pooled over time by district: including and excluding capital goods, weighted and trimmed**

**NOTES**: Estimates are weighted and trimmed using inverse propensity weights that were generated from columns (2) and (5) in Table A1. The first row of figures presents estimates include capital goods, while the second row excludes these from the income measure.



**Figure C3: Five year transition plots pooled over time by district: by type of household, weighted and trimmed**

**NOTES**: Estimates are weighted and trimmed using inverse propensity weights that were generated from columns (2) and (5) in Table A1. The first row of figures presents estimates for households without a wife present, while the second row presents estimates only for households that had wives present.



**Figure C4: Five year transition plots pooled over time and district: by type of production, weighted and trimmed**

**NOTES**: Estimates are weighted and trimmed using inverse propensity weights that were generated from columns (2) and (5) in Table A1. Each figure is limited to income from specific sources.

**Table C1: Mobility estimates for exclusive farmers v. those listed with other occupations in SAF.**