

Supplementary Material for:
Tracking global bicycle ownership patterns

Olufolajimi Oke^{1,2,*}, Kavi Bhalla³, David C. Love^{4,5} and Sauleh Siddiqui^{1,2,6}

¹Department of Civil Engineering, The Johns Hopkins University, Baltimore MD, 21218

²Systems Institute, The Johns Hopkins University, Baltimore, MD 21218

³International Injury Research Unit, The Johns Hopkins Bloomberg School of Public Health, Baltimore MD 21205

⁴Center for a Livable Future, The Johns Hopkins Bloomberg School of Public Health, Baltimore MD 21205

⁵Department of Environmental Health Sciences, The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205

⁶Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, MD 21218

*To whom correspondence should be addressed; E-mail: ooke1@jhu.edu

August 4, 2015

Abstract

In this supplement, we clarify aspects of our extraction and processing of survey data in Section 1. Section 2 contains plots of the bicycle ownership values for all the 150 countries analyzed, while Section 3 includes the complete reference dataset and a dendrogram of the unweighted pair-group method of arithmetic means clustering. All the scripts (written in Python and R) and resources can be found at ce.jhu.edu/sauleh/obls-gbu, and the organizational structure of the code is provided in Section 4.

Contents

1	Data Collection & Methods	3
1.1	Survey mining	3
1.2	Household data	5
2	Percentage bicycle ownership for each country	5
3	Reference data	31
3.1	Processed dataset	31
3.2	Unweighted pair-group method with arithmetic mean: dendrogram	33
4	Code organization	35

List of Figures

S1	Bicycle ownership trends by country	6
S2	Dendrogram of UPGMA clustering	34

List of Tables

S1	Survey sources for bicycle ownership data. *Sources 12, 13 and 14 were unused, as they were educational surveys.	3
S2	Initial country names for which data was available. For analyses, Yugoslavia was considered as Macedonia, West Germany was taken as Germany, and England, Northern Ireland, Scotland and Wales were all considered as the United Kingdom. Data for Seychelles and Zanzibar was only available from the unused educational surveys, hence they do not appear in the final analyses.	4

1 Data Collection & Methods

In this section, we describe how household bicycle availability and household population data was collected and processed from sources listed.

1.1 Survey mining

Initially, we obtained bicycle ownership data from the surveys shown in [Supplementary Table S1](#) conducted in various countries over 1970–2012. We used 1989 as the starting year for our analyses, as data for preceding years was only available for Chile, Indonesia, Malaysia and Thailand. In these surveys, the relevant inquiry determined in a binary fashion if the households had one bicycle (or more) available. Many survey records were accessed using Stata and then compiled into a spreadsheet. Appendix A contains ownership data plots for all the countries, indicating the survey source of each point.

Supplementary Table S1 Survey sources for bicycle ownership data. *Sources 12, 13 and 14 were unused, as they were educational surveys.

ID	Source	Acronym	Years conducted
1	World Health Surveys [1]	WHS	2002
2	Demographic and Health Surveys [2]	DHS	Multiple
3	Malaria Indicator Surveys [3]	MIS	2006-7, 2008-9
4	Integrated Public Use Microdata Services [4]	IPUMS	Multiple
5	International Crime Victim Surveys [5]	ICVS	1989-2002
6	India National Census [6]	INC	2001, 2011
7	Multiple Indicator Cluster Surveys 4 [7]	MICS4	2010-2011
8	Multiple Indicator Cluster Surveys 3 [8]	MICS3	2005-2009
9	Study on Global Ageing and Adult Health [9]	SAGE	2007-2011
10	Integrated Survey on the Welfare of the Population [10]	IBEP	2008-2009
11	Enquete Demographique et de Sante et a Indicateurs Multiples [11]	EDSM	2006
*12	Southern & Eastern Africa Consortium for Monitoring Educational Quality [12]	SACMEQ3	2005
*13	Southern and Eastern Africa Consortium for Monitoring Educational Quality [13]	SACMEQ2	1999
*14	Southern and Eastern Africa Consortium for Monitoring Educational Quality [13]	SACMEQ1	1995

The educational surveys conducted by the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) were found to not be indicative of the national populations in comparison with data from other surveys, since the primary respondents were school pupils. The SACMEQ results were therefore not included in our analyses. For some countries, ownership data was available from multiple surveys within the same year. In initializing the data, we took the average value of the country-years where this was the case (for example, Vietnam had data available from both WHS and DHS in 2002). Also, data from any of the countries in the United Kingdom (England, Northern Ireland, Scotland and Wales) was taken as representative of the UK. Finally, the International Crime Victim Surveys (ICVS) 1996 datapoint for Yugoslavia was considered valid for Macedonia. We therefore ended up with 150 countries after this pre-processing step.

There are a total of 452 data points. [Supplementary Table S2](#) lists the initial countries in our dataset and their corresponding three-letter acronyms (ISO) before pre-processing.

Supplementary Table S2 Initial country names for which data was available. For analyses, Yugoslavia was considered as Macedonia, West Germany was taken as Germany, and England, Northern Ireland, Scotland and Wales were all considered as the United Kingdom. Data for Seychelles and Zanzibar was only available from the unused educational surveys, hence they do not appear in the final analyses.

Country	ISO	Country	ISO	Country	ISO
Afghanistan	AFG	Greece	GRC	Peru	PER
Albania	ALB	Guatemala	GTM	Philippines	PHL
Angola	AGO	Guinea	GIN	Poland	POL
Argentina	ARG	Guinea-Bissau	GNB	Portugal	PRT
Armenia	ARM	Guyana	GUY	Republic of Korea	KOR
Australia	AUS	Haiti	HTI	Republic of Moldova	MDA
Austria	AUT	Honduras	HND	Romania	ROM
Azerbaijan	AZE	Hungary	HUN	Russia	RUS
Bangladesh	BGD	India	IND	Rwanda	RWA
Belarus	BLR	Indonesia	IDN	Sao Tome and Principe	STP
Belgium	BEL	Iraq	IRQ	Scotland	XSC
Belize	BLZ	Ireland	IRL	Senegal	SEN
Benin	BEN	Israel	ISR	Serbia	ZZZX
Bhutan	BTN	Italy	ITA	Seychelles	SYC
Bolivia	BOL	Japan	JPN	Sierra Leone	SLE
Bosnia and Herzegovina	BIH	Jordan	JOR	Slovakia	SVK
Botswana	BWA	Kazakhstan	KAZ	Slovenia	SVN
Brazil	BRA	Kenya	KEN	Somalia	SOM
Bulgaria	BGR	Kyrgyzstan	KGZ	South Africa	ZAF
Burkina Faso	BFA	Laos	LAO	South Sudan	SSD
Burundi	BDI	Latvia	LVA	Spain	ESP
Cambodia	KHM	Lebanon	LBN	Sri Lanka	LKA
Cameroon	CMR	Lesotho	LSO	Sudan	SDN
Canada	CAN	Liberia	LBR	Suriname	SUR
Central African Republic	CAF	Lithuania	LTU	Swaziland	SWZ
Chad	TCD	Luxembourg	LUX	Sweden	SWE
Chile	CHL	Macedonia	MKD	Switzerland	CHE
China	CHN	Madagascar	MDG	Tajikistan	TJK
Colombia	COL	Malawi	MWI	Tanzania	TZA
Comoros	COM	Malaysia	MYS	Thailand	THA
Congo	COG	Maldives	MDV	Timor-Leste	TLS
Congo DRC	COD	Mali	MLI	Togo	TGO
Costa Rica	CRI	Malta	MLT	Trinidad and Tobago	TTO
Cote d'Ivoire	CIV	Mauritania	MRT	Tunisia	TUN
Croatia	HRV	Mauritius	MUS	Turkey	TUR
Czech Republic	CZE	Mexico	MEX	Turkmenistan	TKM
Denmark	DNK	Mongolia	MNG	Uganda	UGA
Djibouti	DJI	Montenegro	MNE	Ukraine	UKR
Dominican Republic	DOM	Morocco	MAR	United Arab Emirates	ARE
Ecuador	ECU	Mozambique	MOZ	United Kingdom England and Wales	XWA
Egypt	EGY	Myanmar	MMR	United Kingdom Northern Ireland	XNI
England	XWA	Namibia	NAM	United Kingdom Scotland	XSC
Eritrea	ERI	Nepal	NPL	United States	USA
Estonia	EST	Netherlands	NLD	Uruguay	URY
Ethiopia	ETH	New Zealand	NZL	Uzbekistan	UZB
Finland	FIN	Nicaragua	NIC	Vanuatu	VUT
France	FRA	Niger	NER	Venezuela	VEN
Gabon	GAB	Nigeria	NGA	Vietnam	VNM
Gambia	GMB	Northern Ireland	XNI	Yemen	YEM
Georgia	GEO	Norway	NOR	Yugoslavia	YUG
Germany	DEU	Pakistan	PAK	Zambia	ZMB
Germany West	ZZZY	Panama	PAN	Zanzibar	ZZZU
Ghana	GHA	Paraguay	PRY	Zimbabwe	ZWE

1.2 Household data

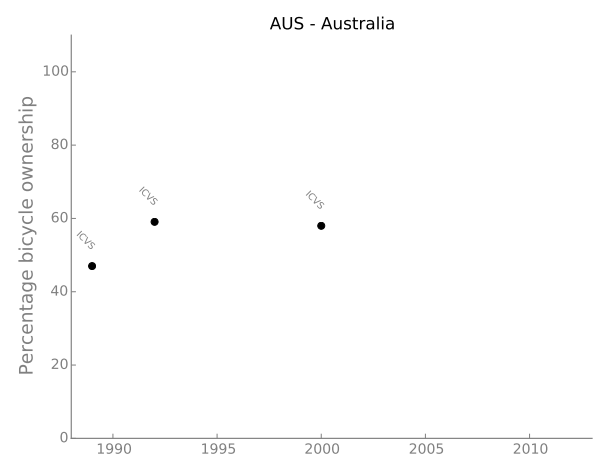
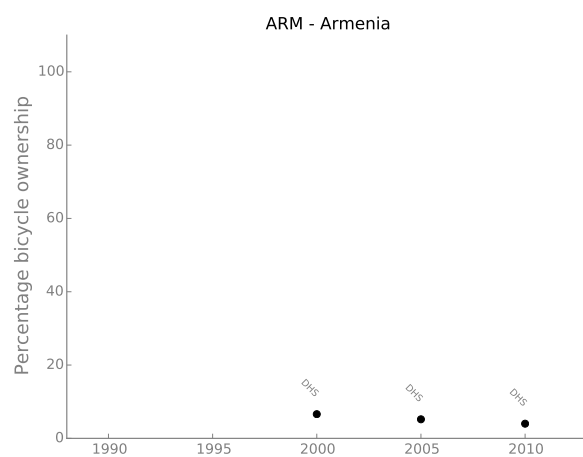
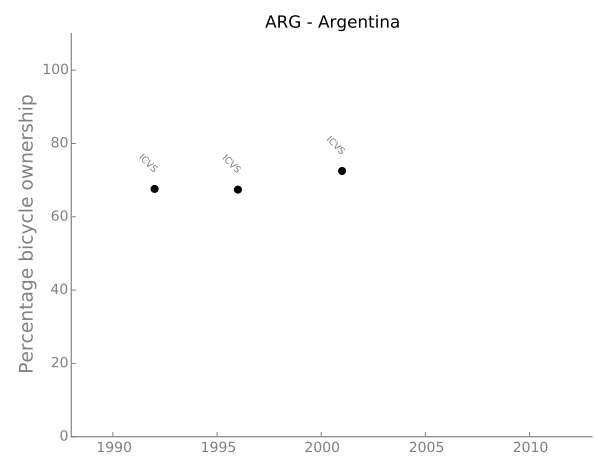
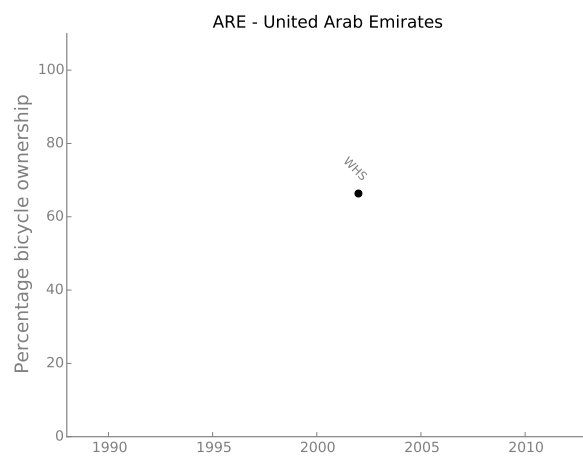
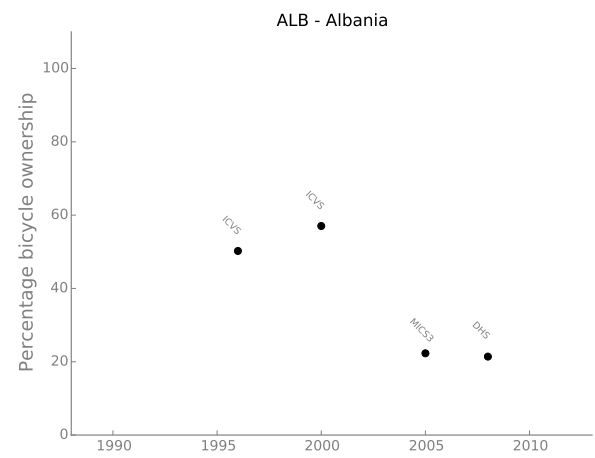
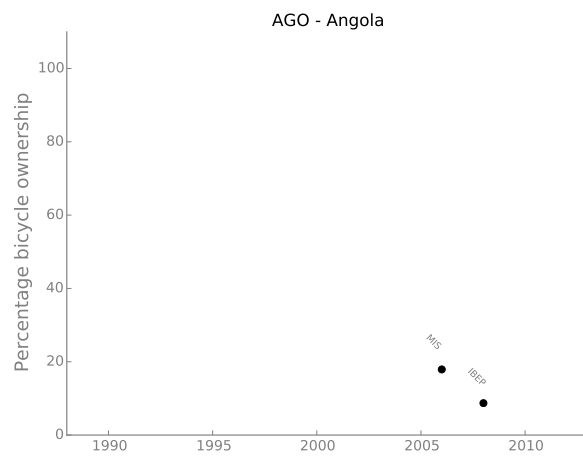
The Integrated Public Use Microdata Services [4] and World Health Surveys [1] records also contained estimates of the number of households in respondent countries, but for countries in other surveys we had to obtain these numbers from the United Nations [14], the United Nations Economic Commission for Europe (UNECE) [15] and other miscellaneous sources. We only searched for household populations in the years for which bicycle ownership data were available. In cases where household population numbers were not available, we estimated the number of households by multiplying the average household size by the population of the country (obtained from World Bank databases [16]). For countries where household numbers were not available in certain years, we used data for the closest earlier year.

The processing of household data is summarized in the following steps:

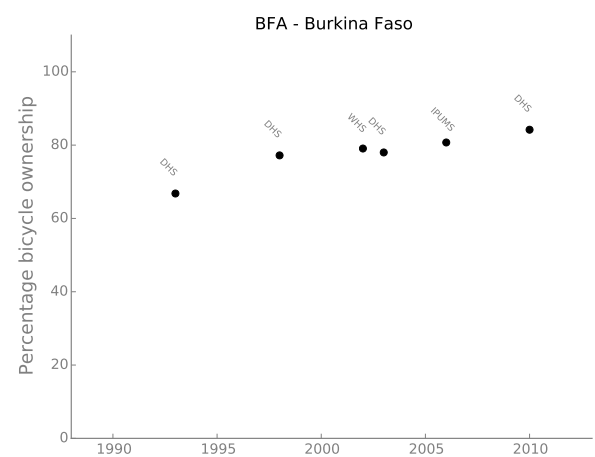
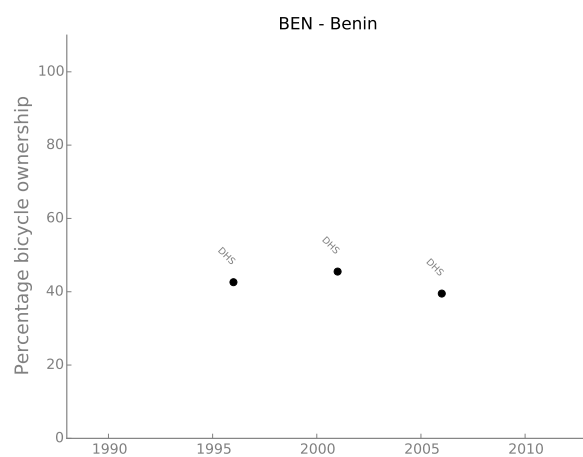
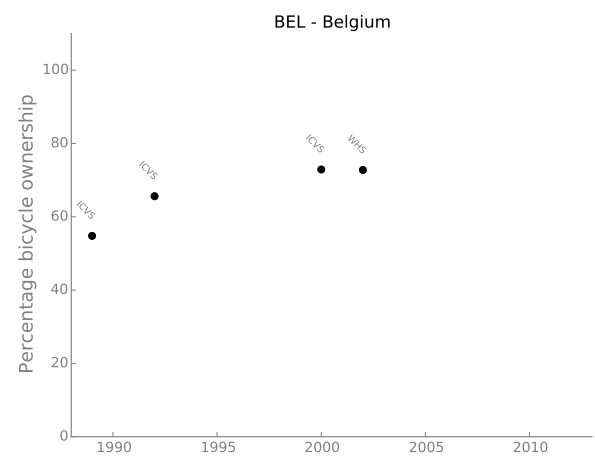
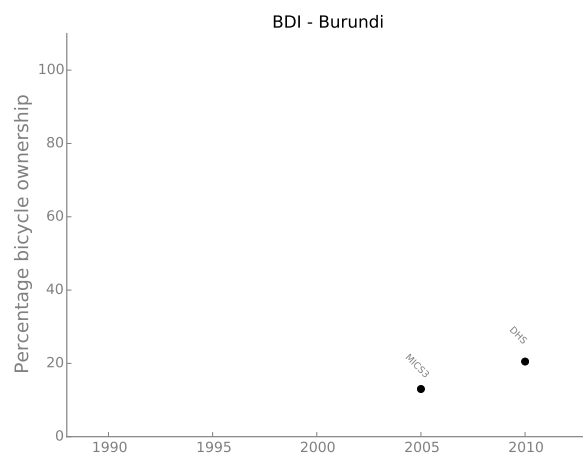
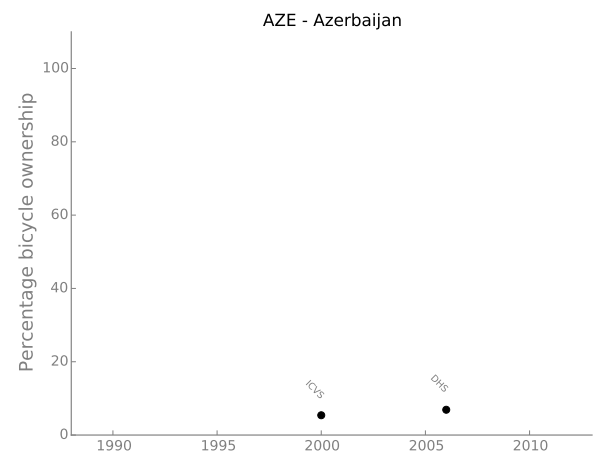
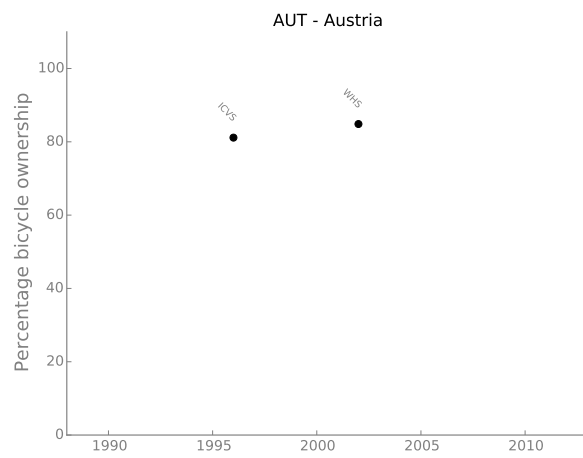
1. Search for number of households for countries by year.
2. If household data are not available for a country in a given year, use data for the closest earlier year.
3. If, for a certain year, step 2. is infeasible within a 20-year span, use the maximum household population available for that country in any year
4. If the number of households for a certain country is completely missing for a required year, take the average household size, and multiply by its population in that year to estimate the number of households.

2 Percentage bicycle ownership for each country

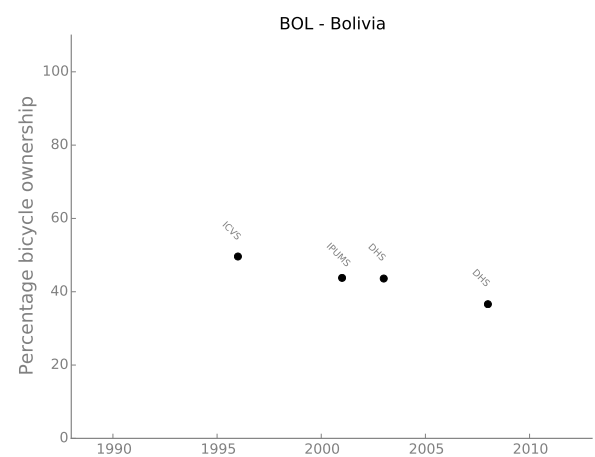
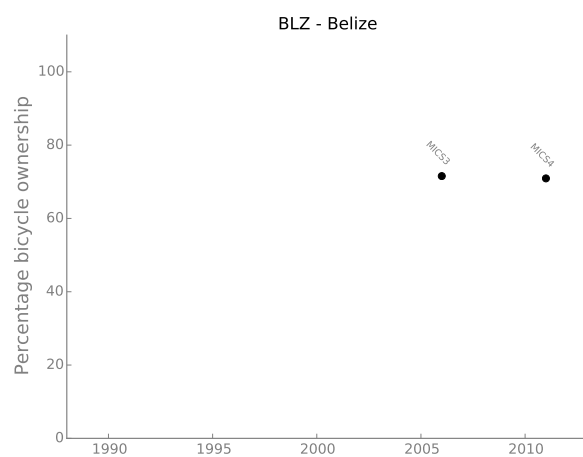
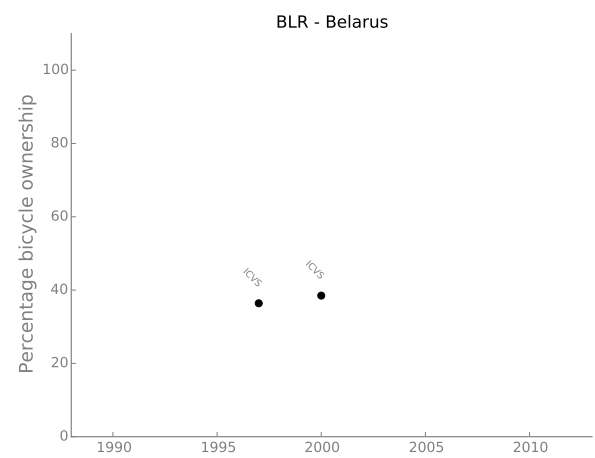
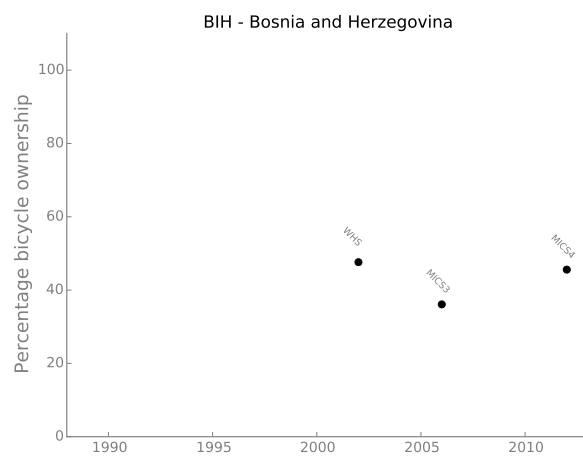
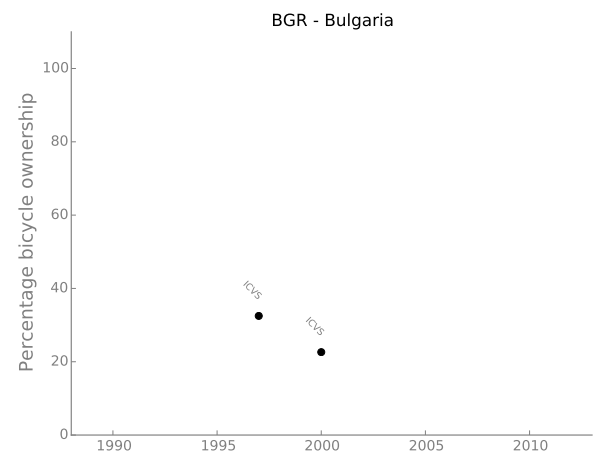
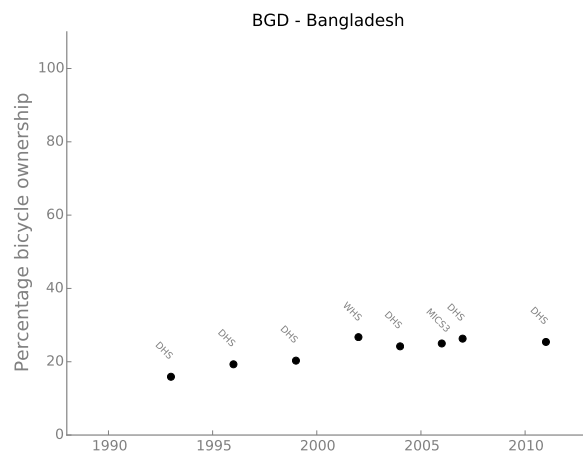
[Supplementary Fig. S1](#) shows individual bicycle ownership plots for each of the countries studied. The survey sources have also been labeled (refer to [Supplementary Table S1](#)). For all instances in which more than one datapoint was available for a country-year (e.g. Dominican Republic-2002), we used an average in the pre-processing stage, but multiple country-years are shown in these plots to indicate survey diversity.



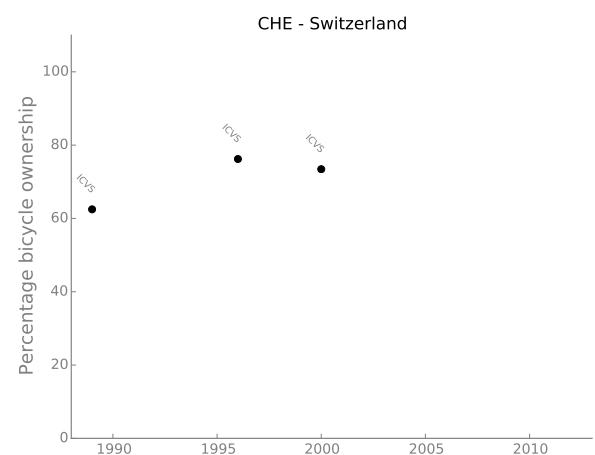
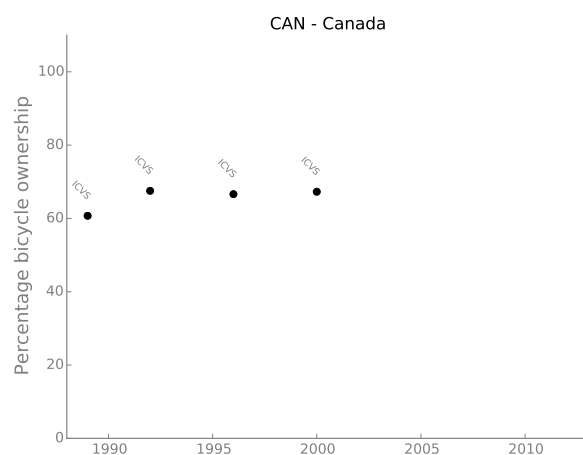
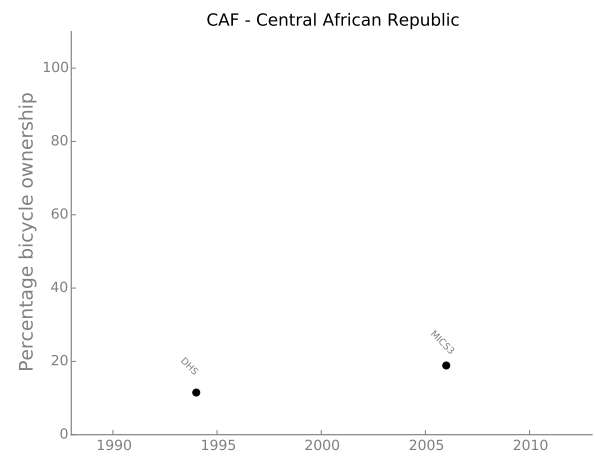
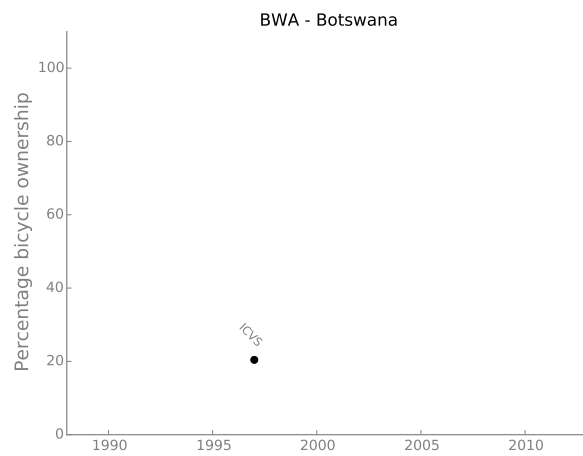
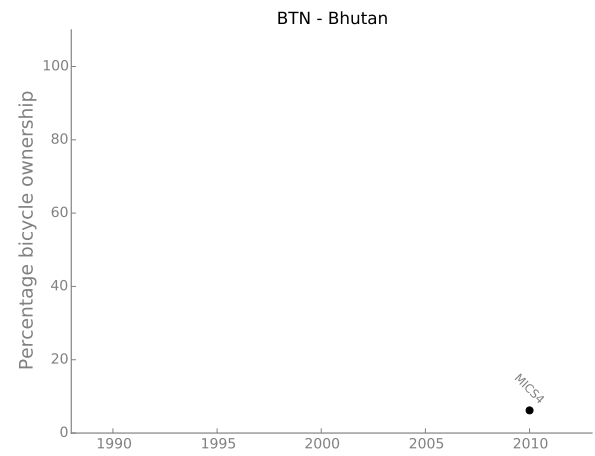
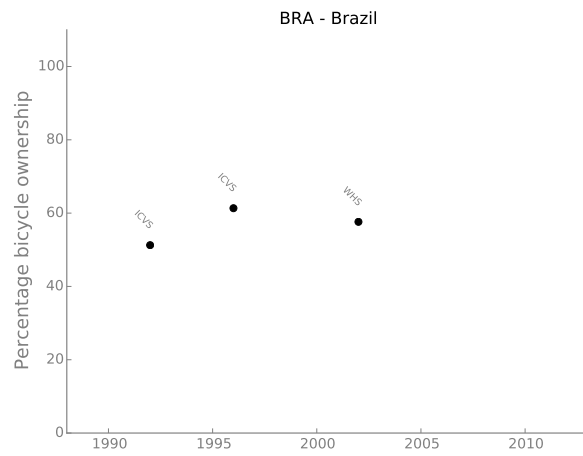
Supplementary Fig. S1 Bicycle ownership trends by country



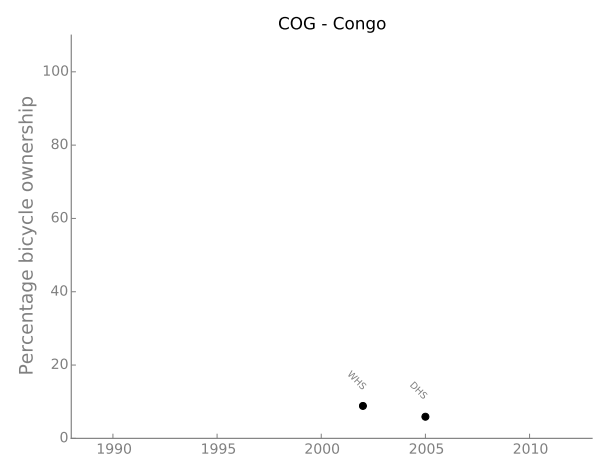
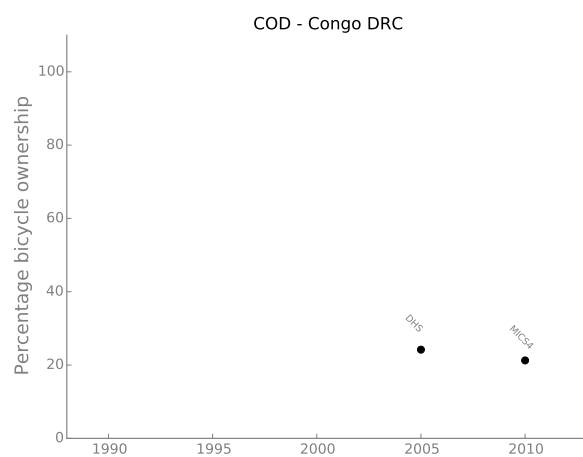
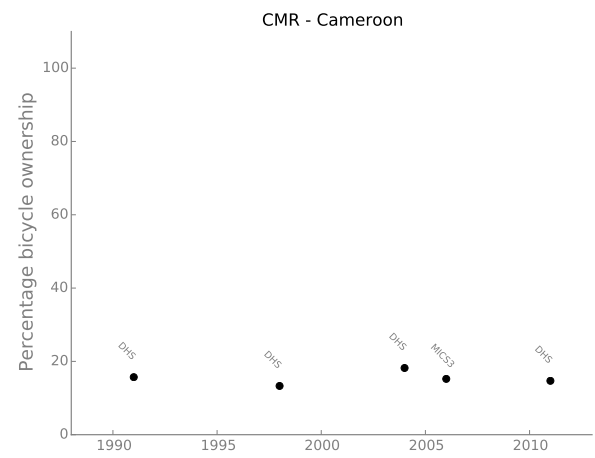
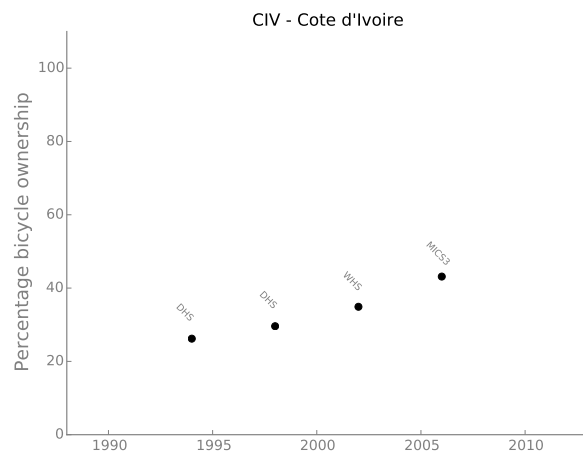
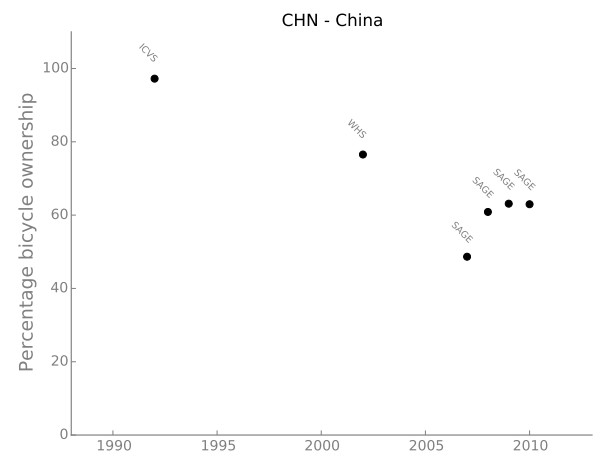
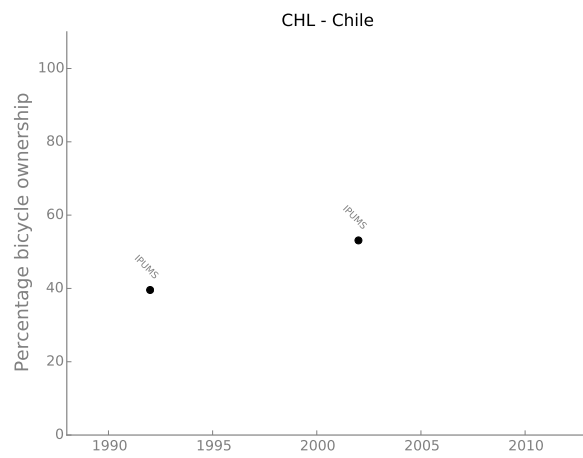
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



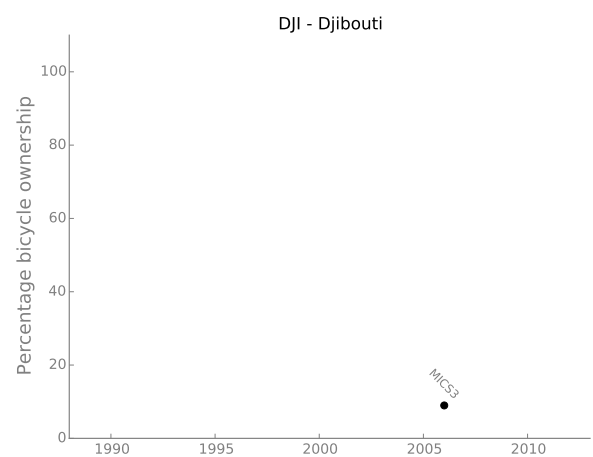
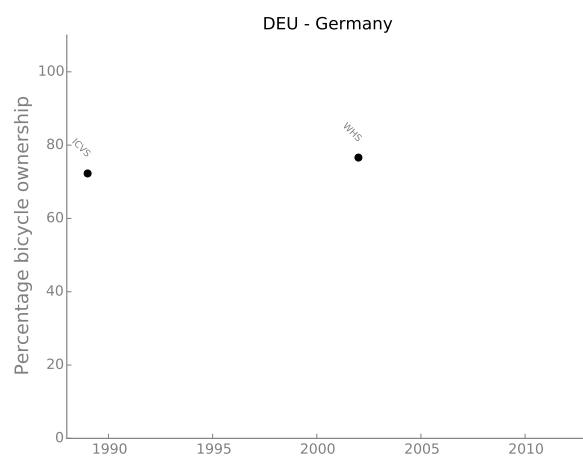
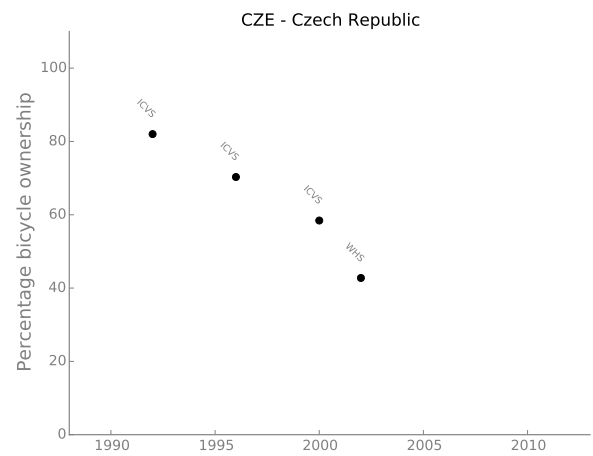
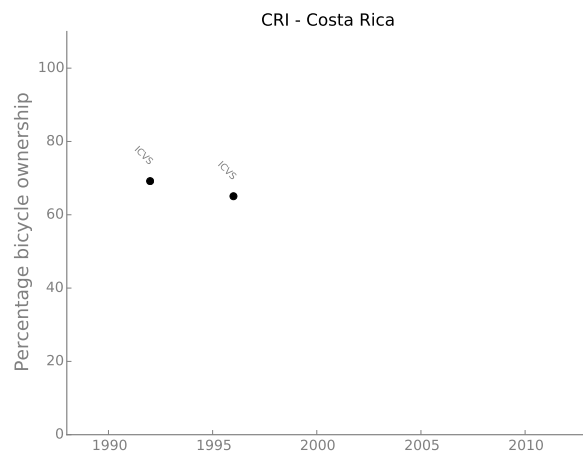
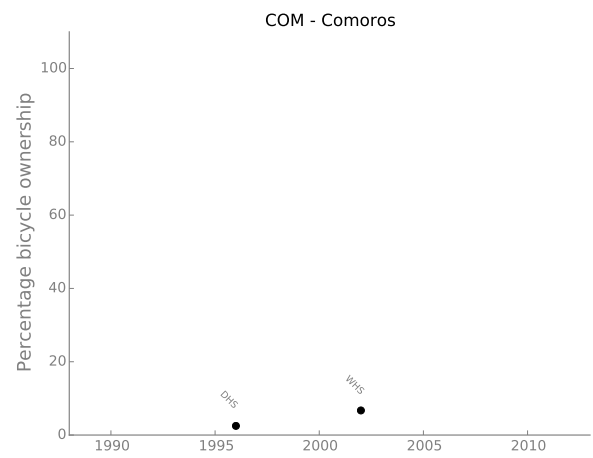
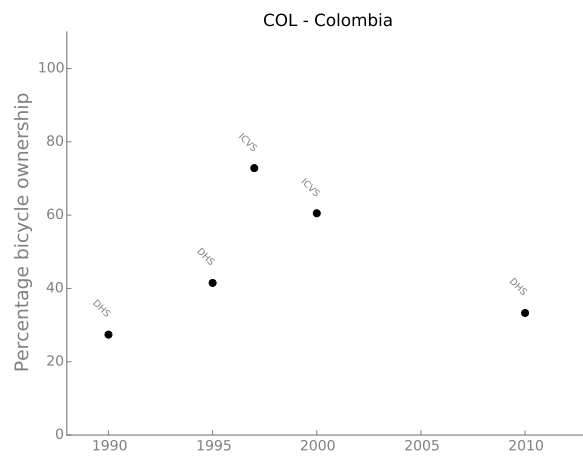
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



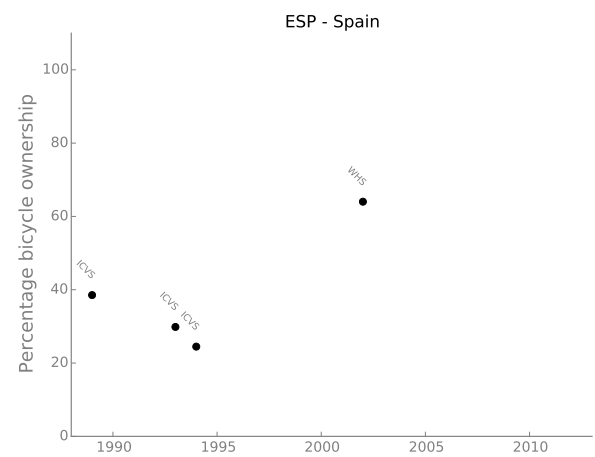
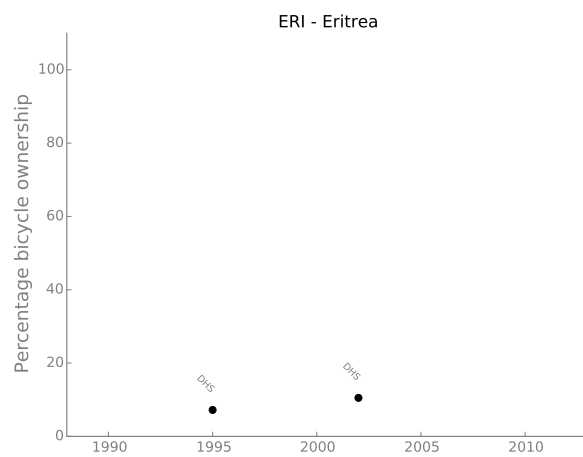
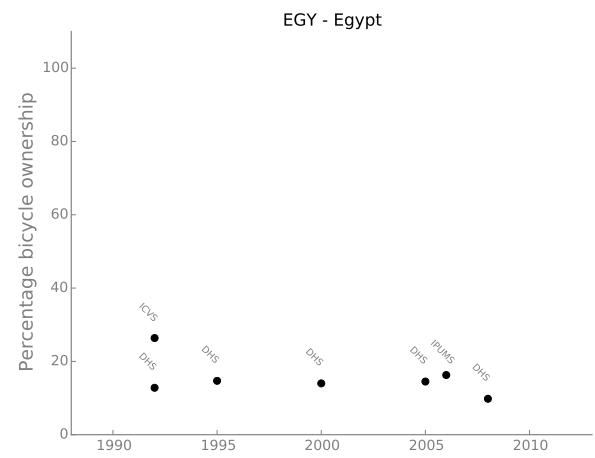
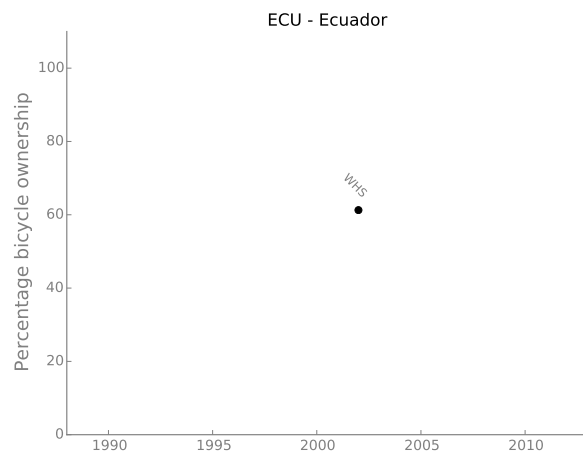
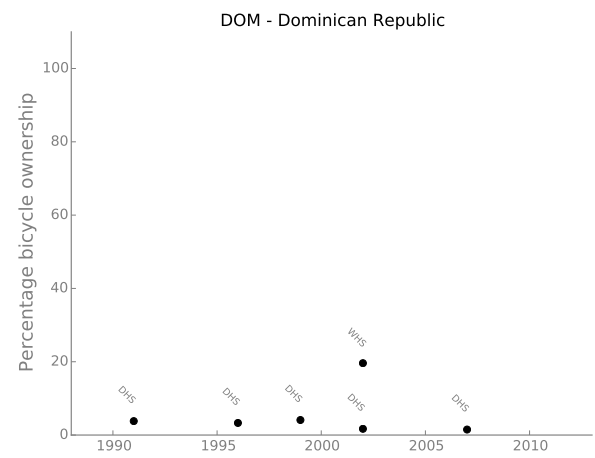
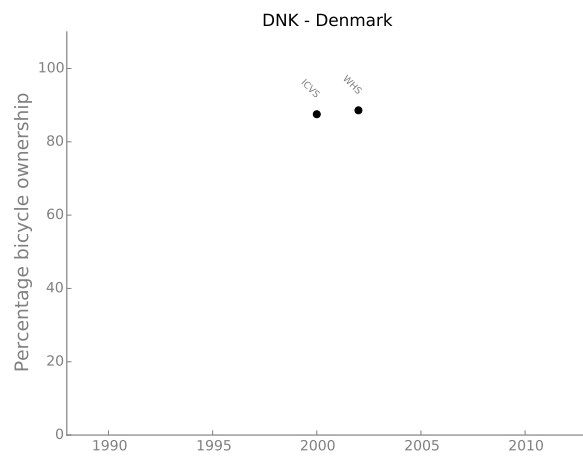
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



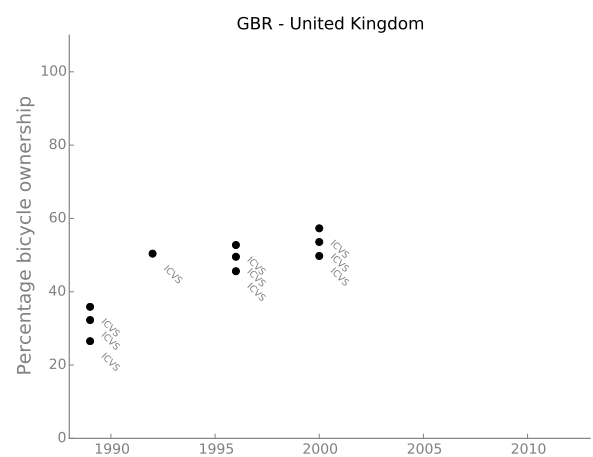
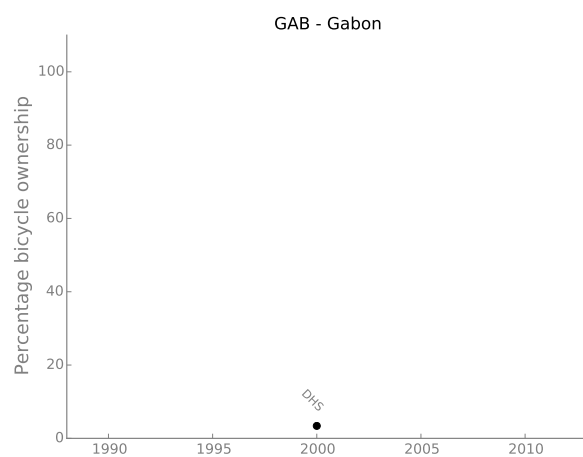
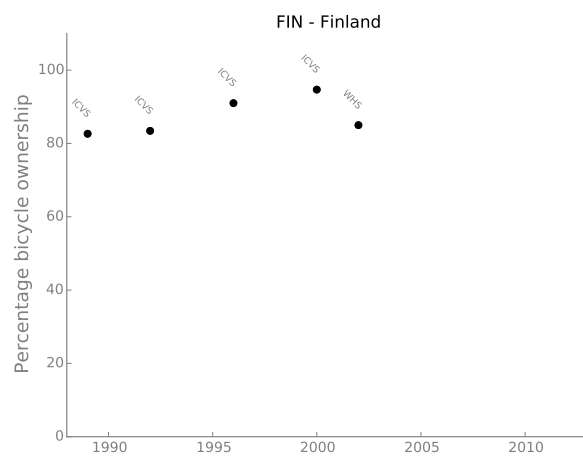
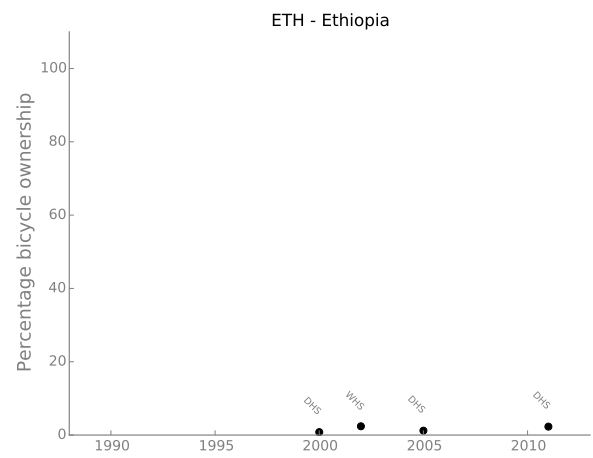
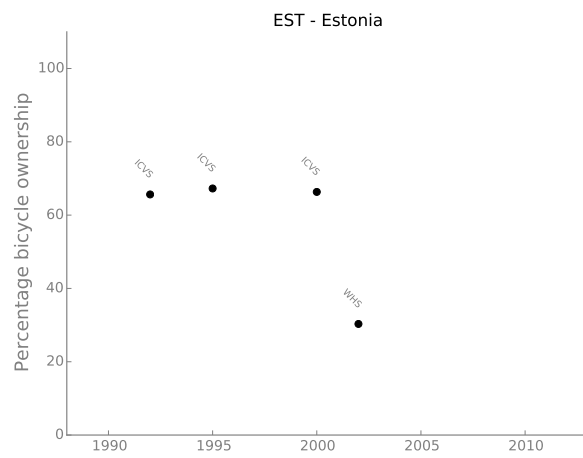
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



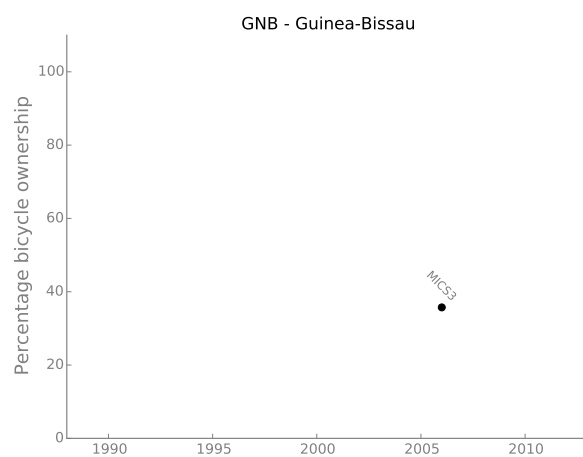
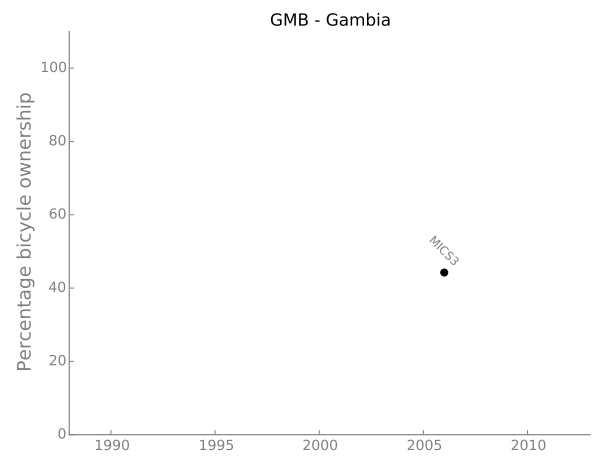
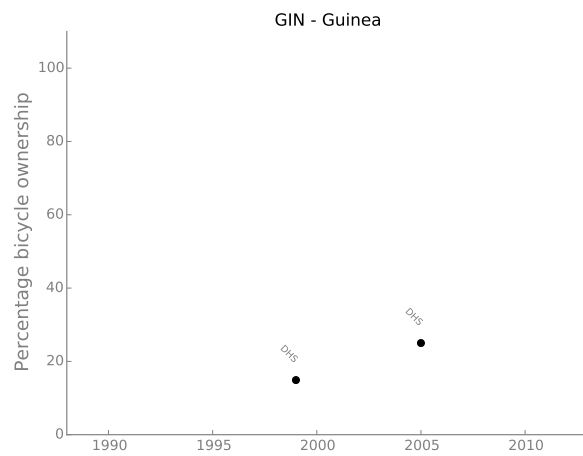
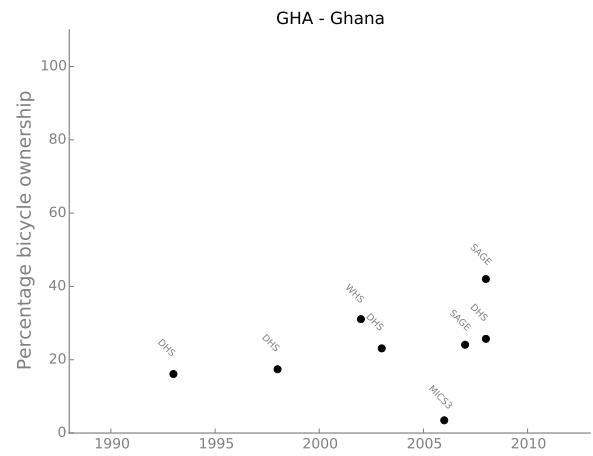
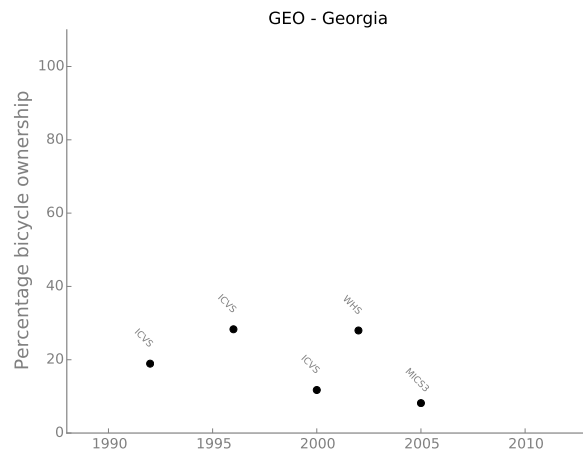
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



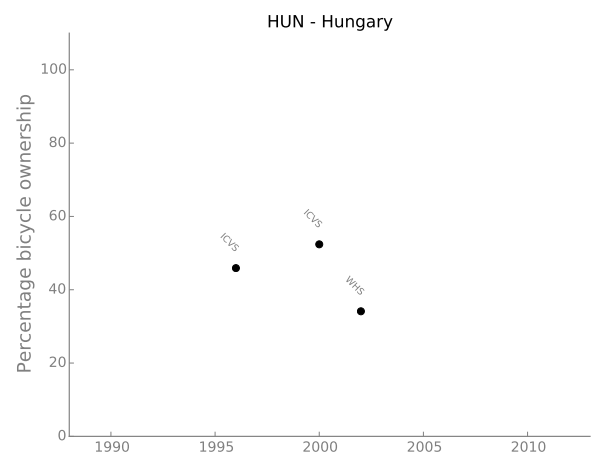
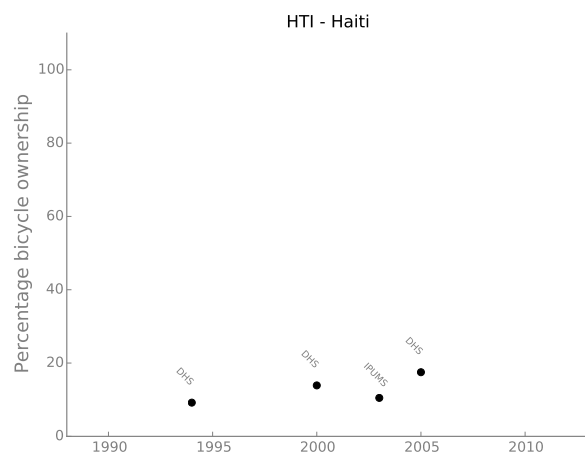
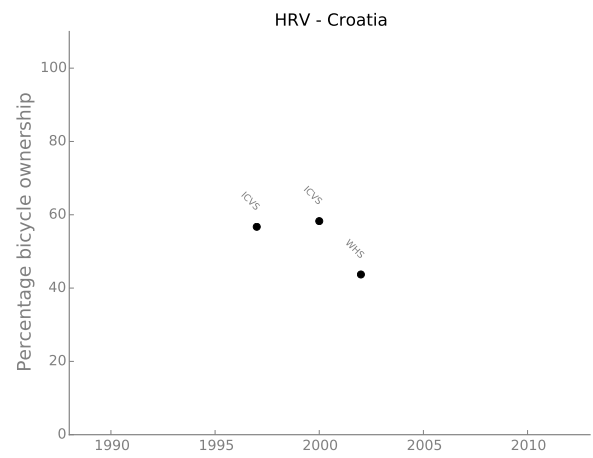
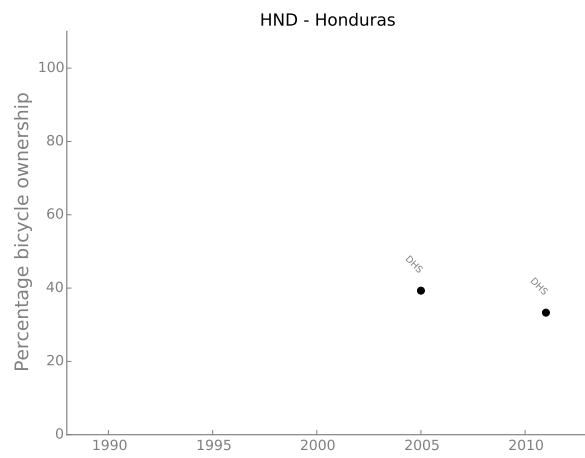
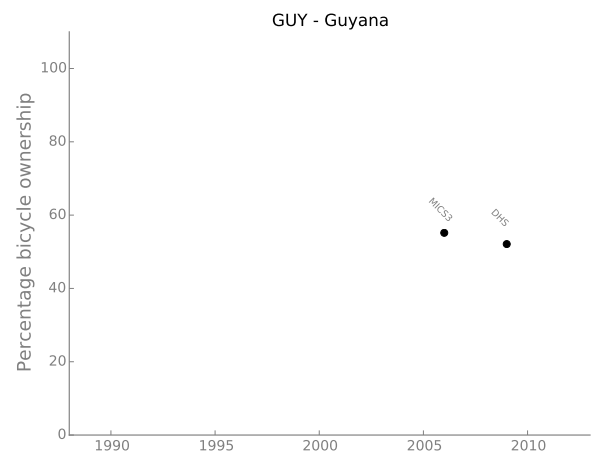
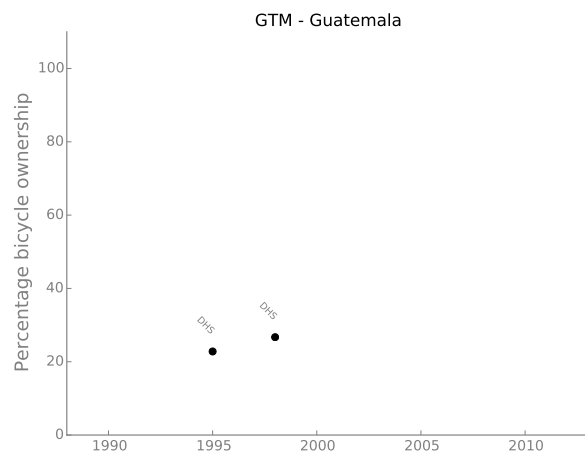
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



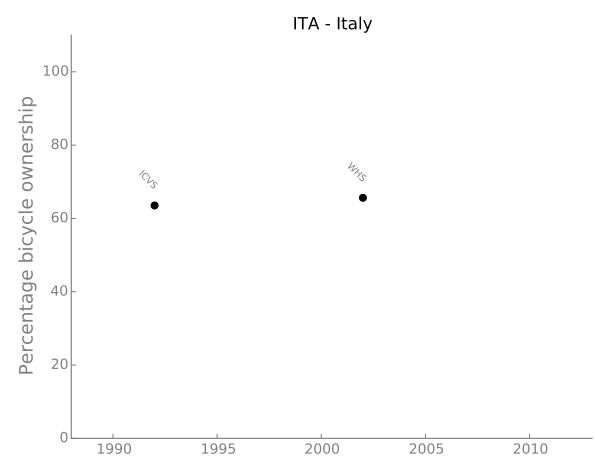
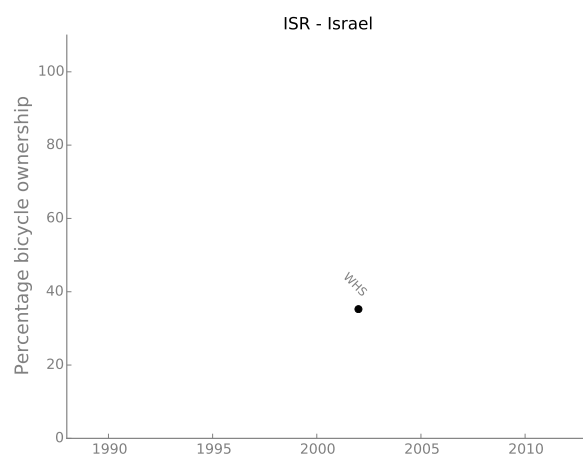
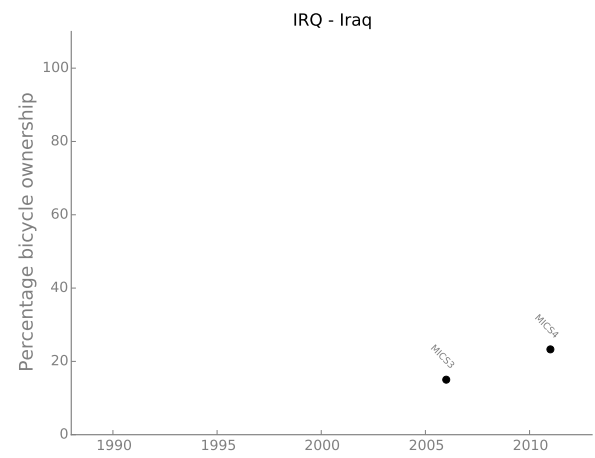
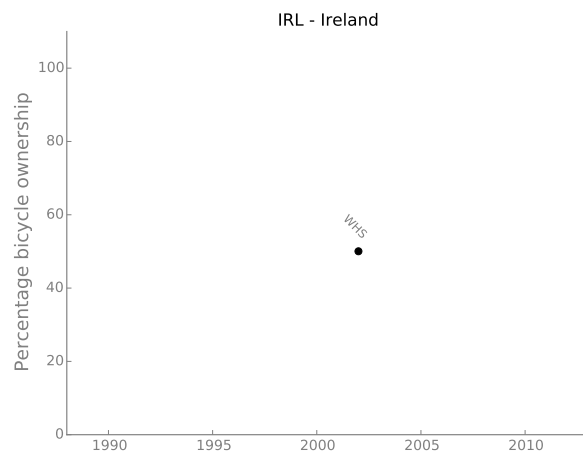
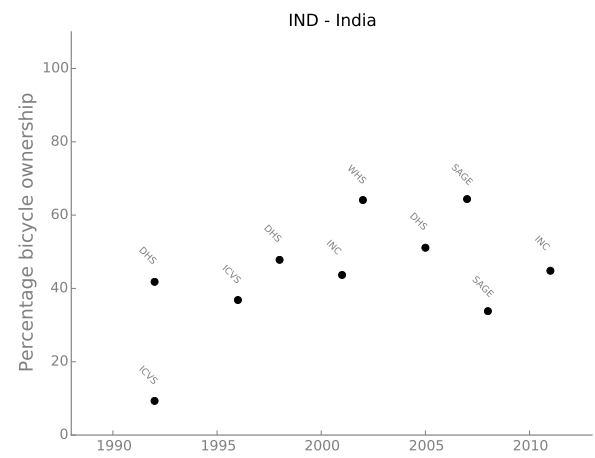
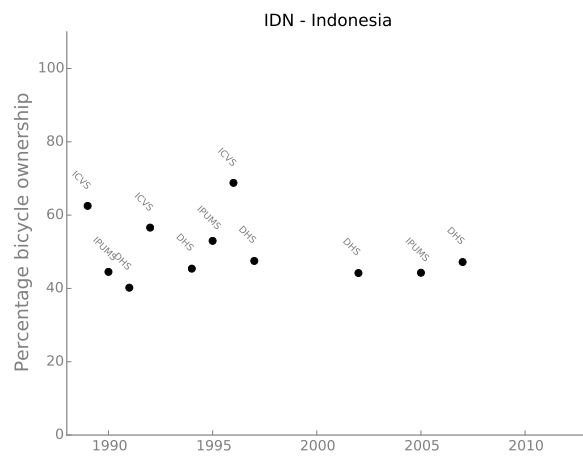
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



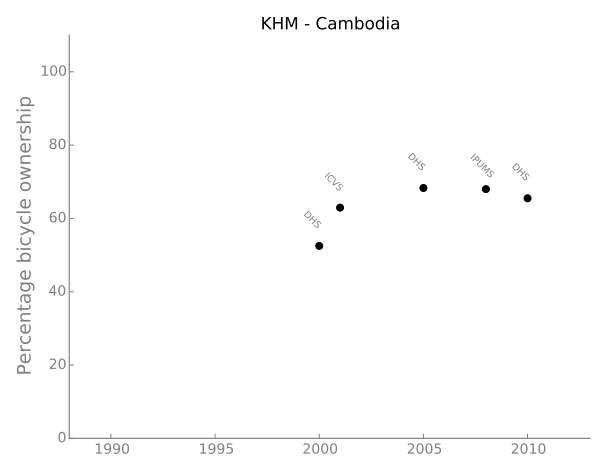
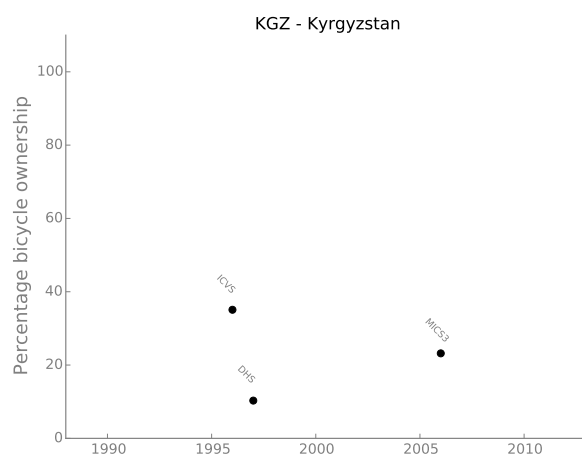
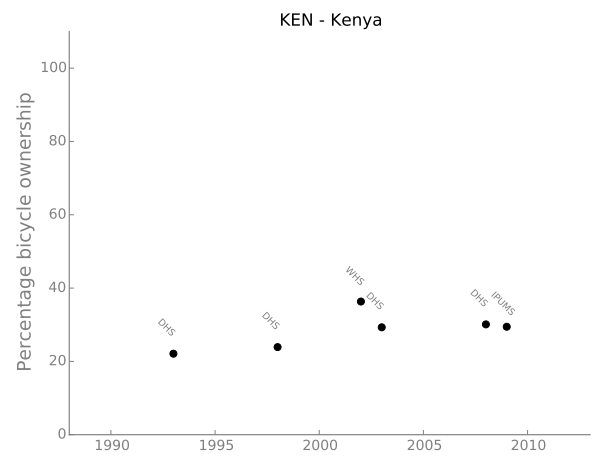
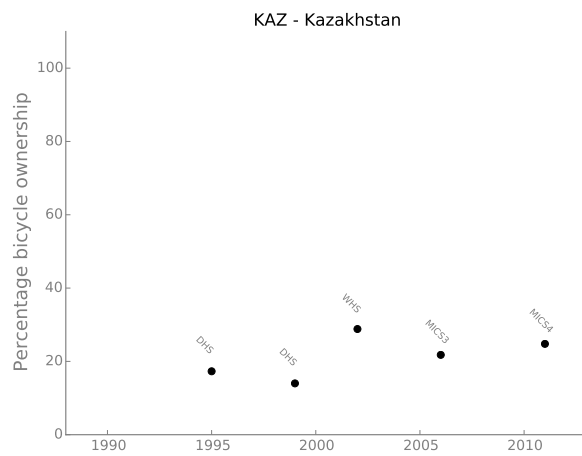
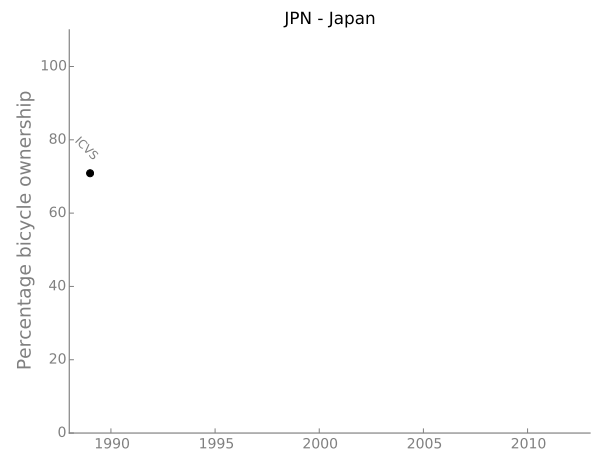
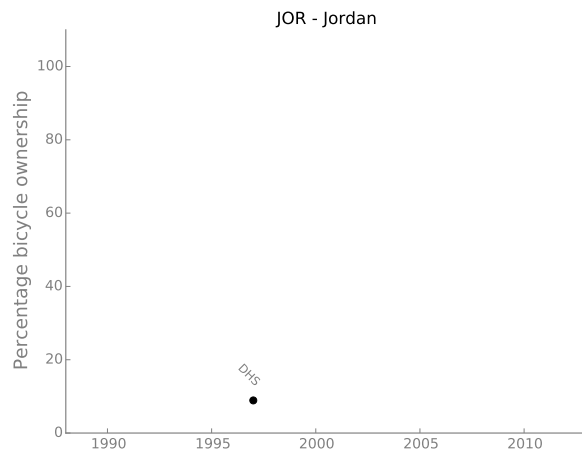
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



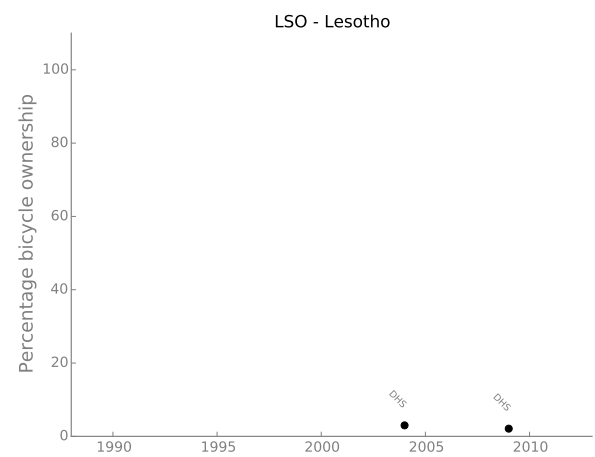
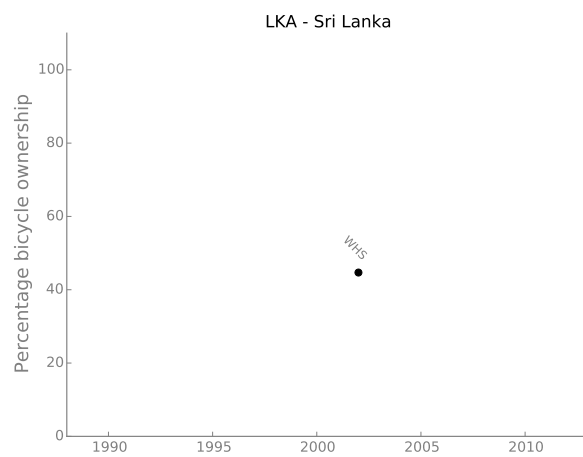
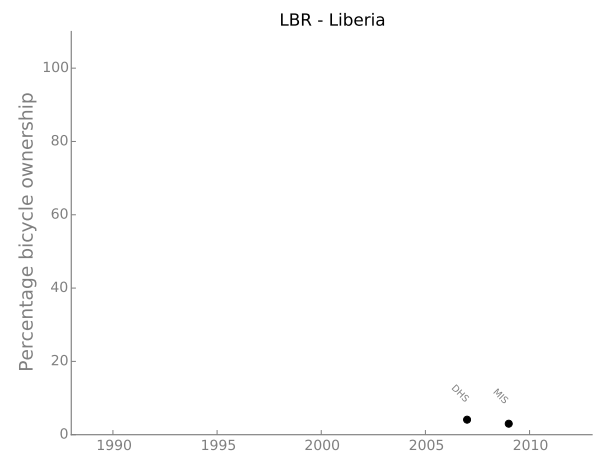
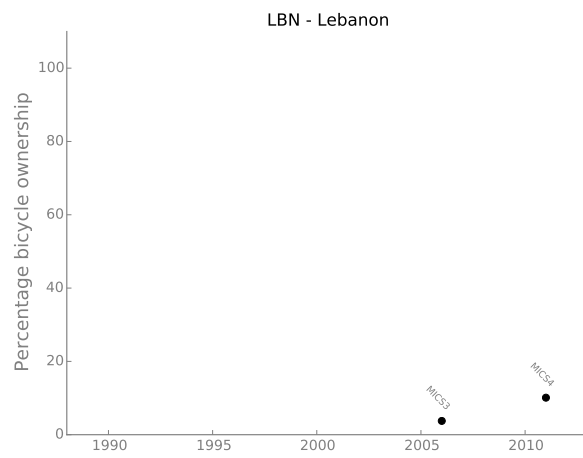
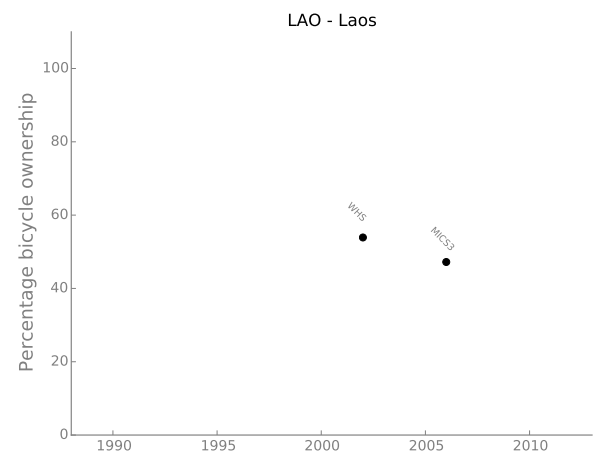
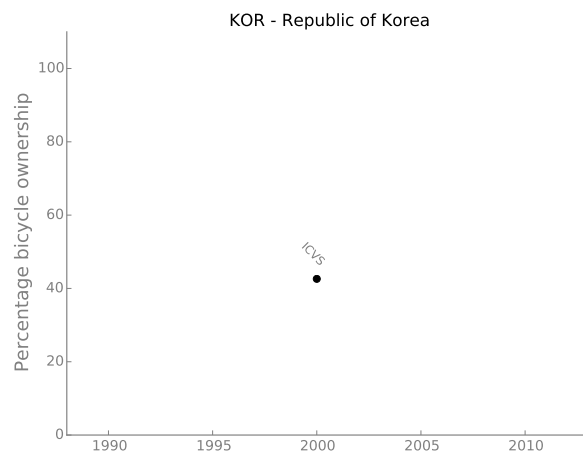
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



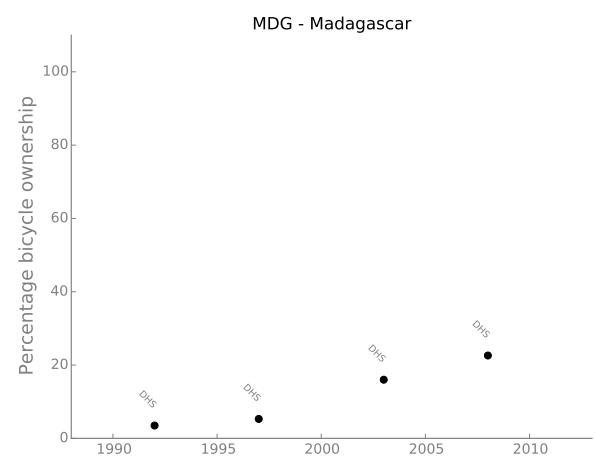
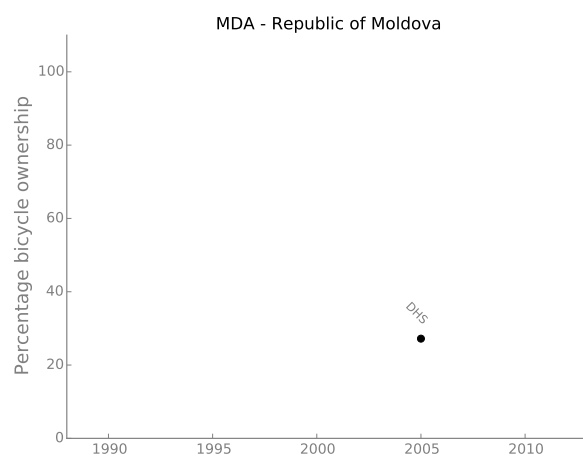
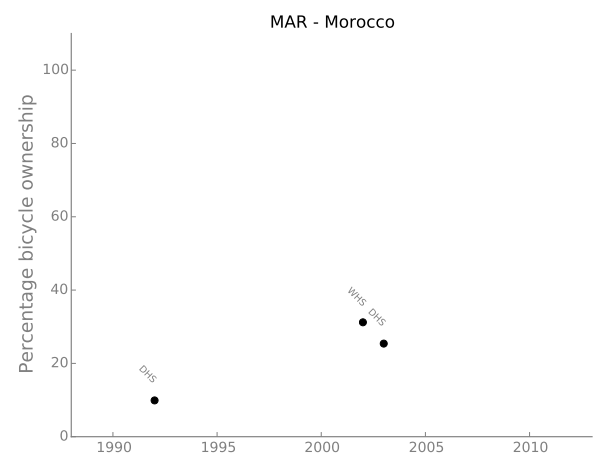
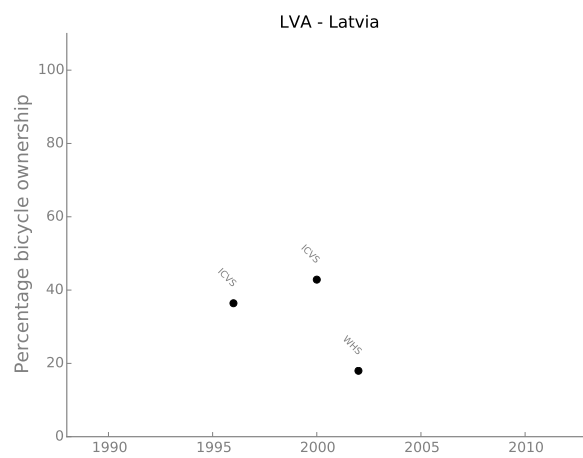
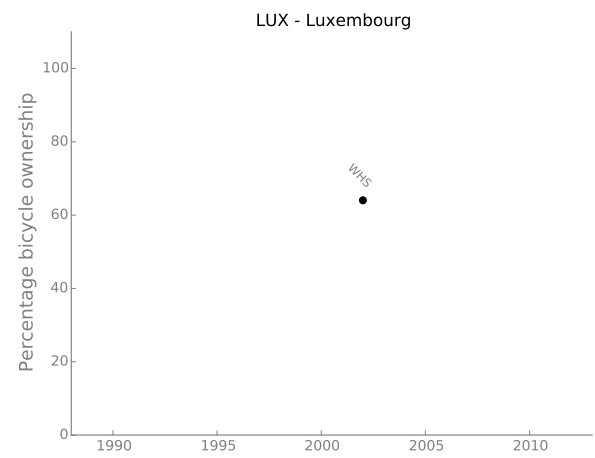
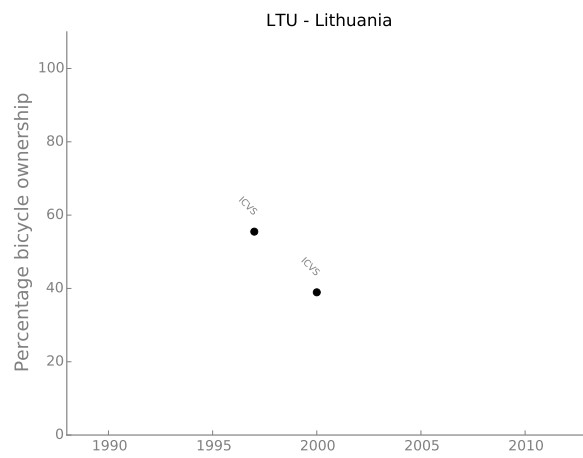
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



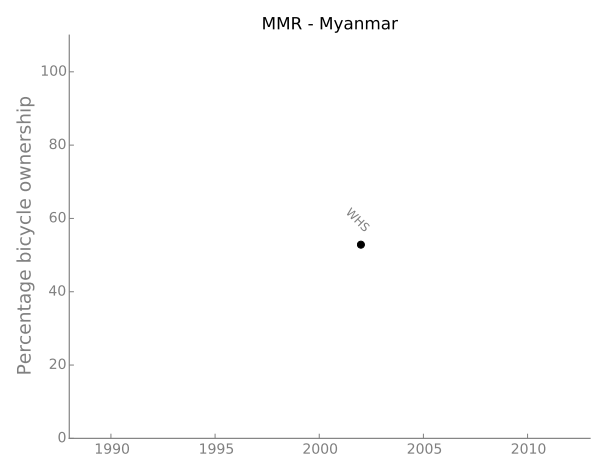
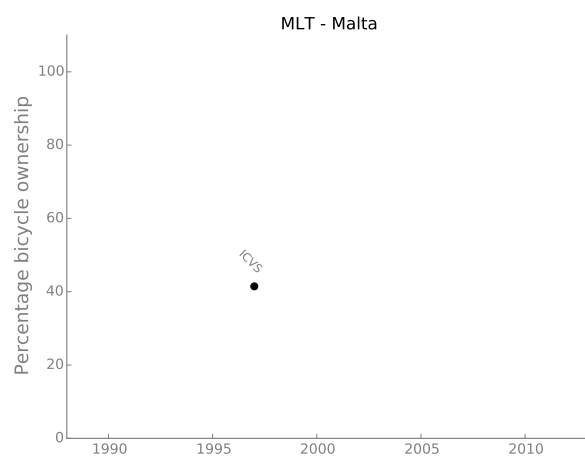
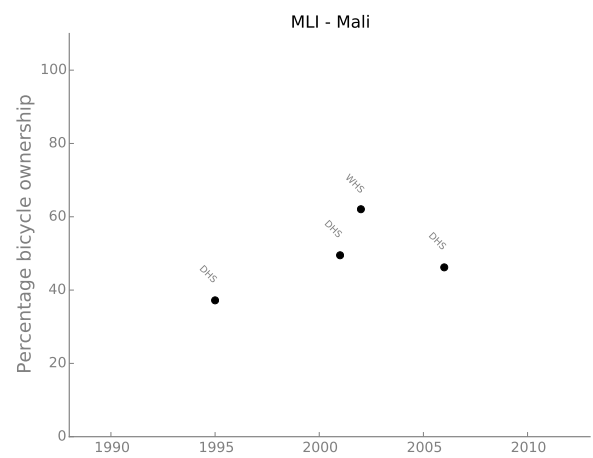
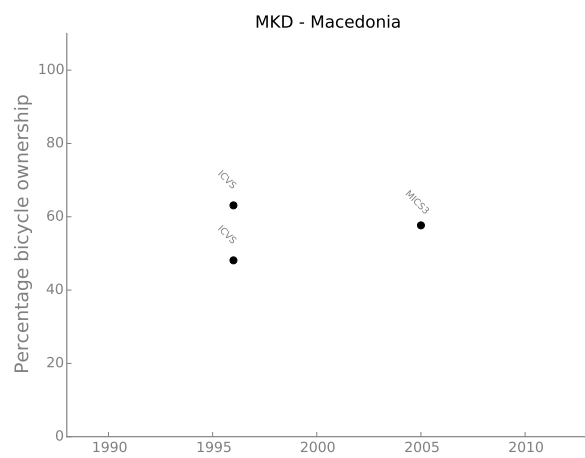
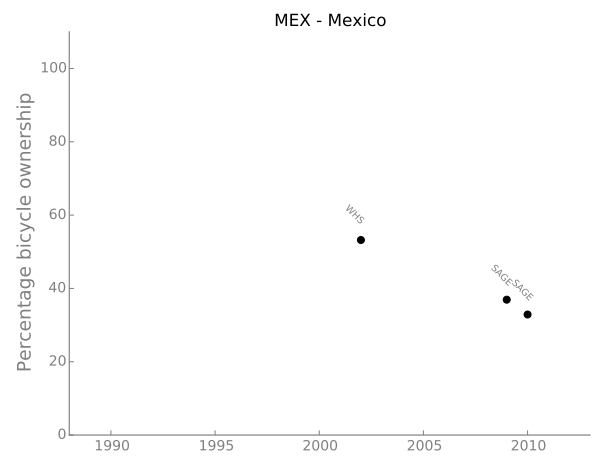
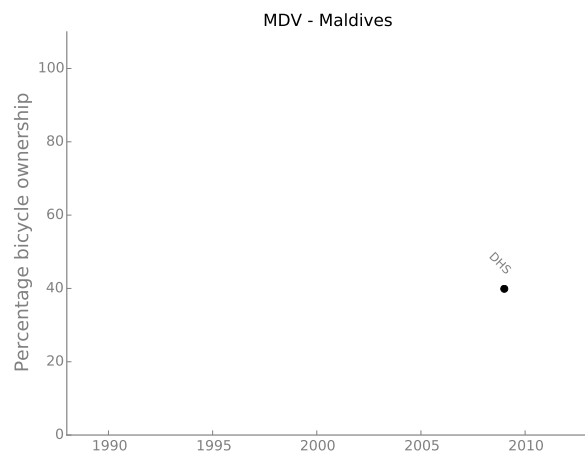
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



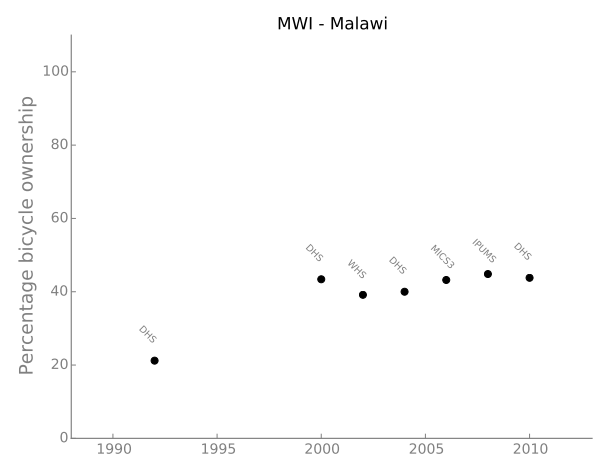
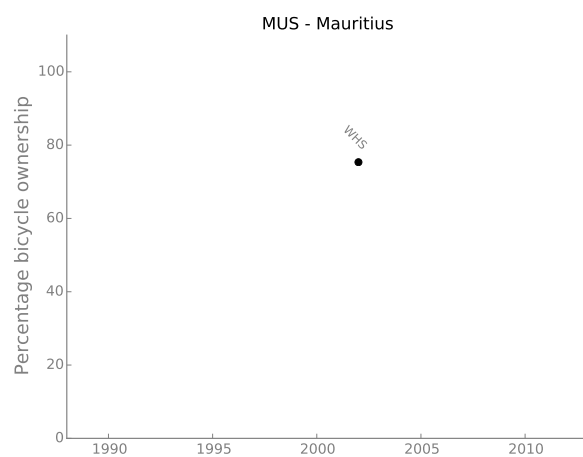
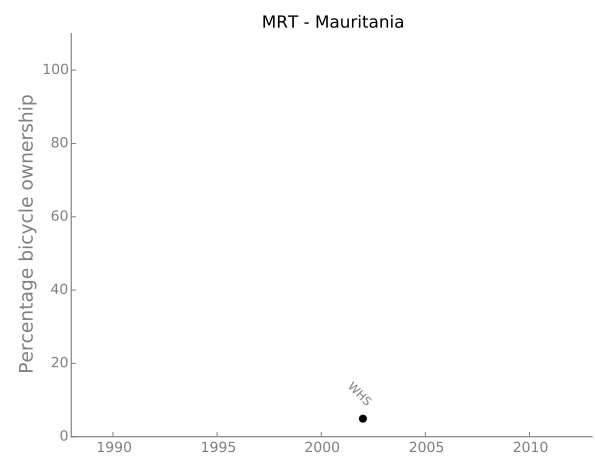
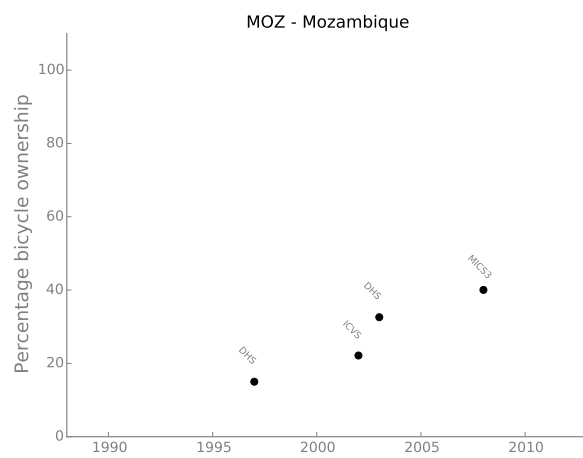
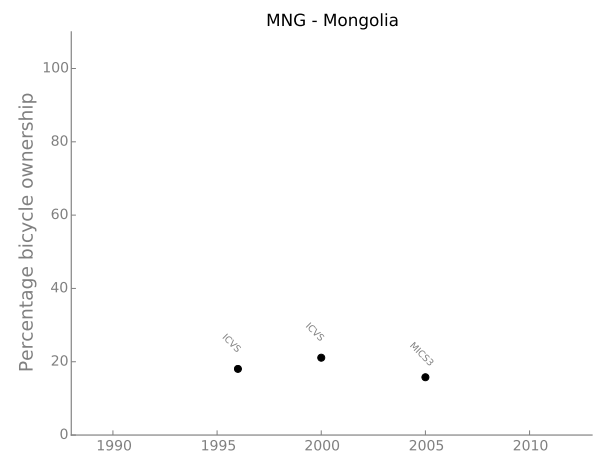
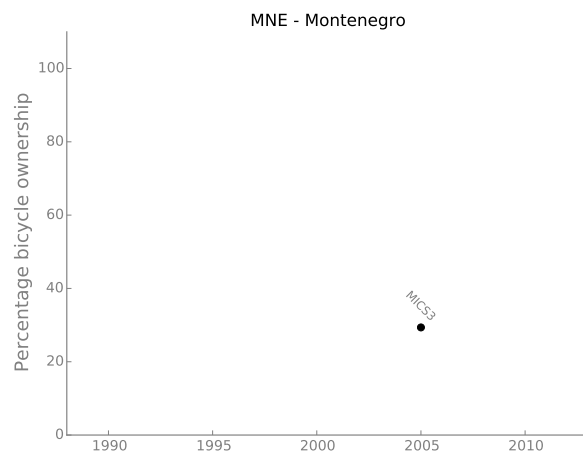
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



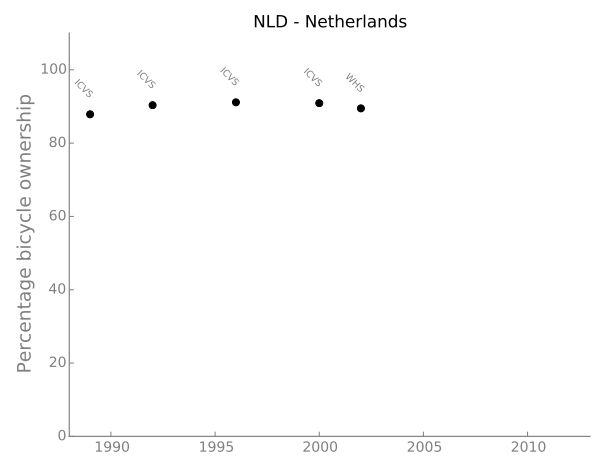
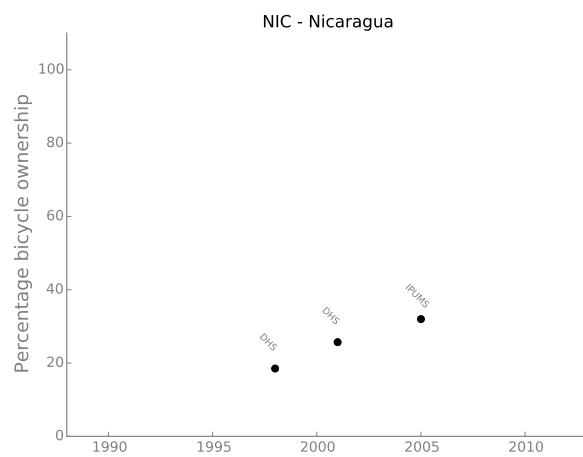
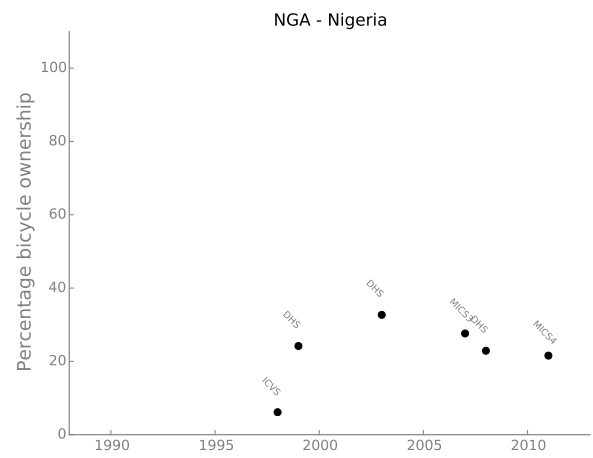
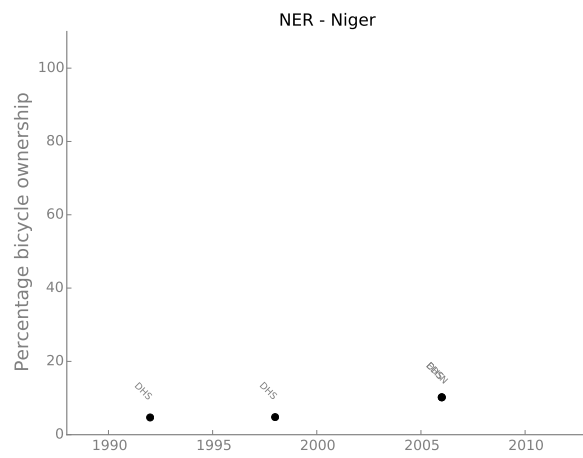
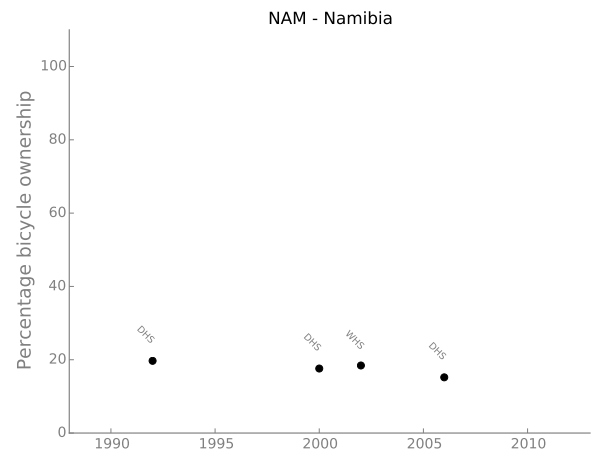
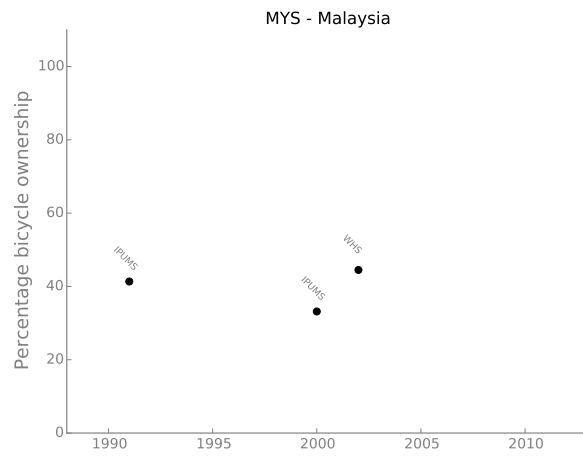
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



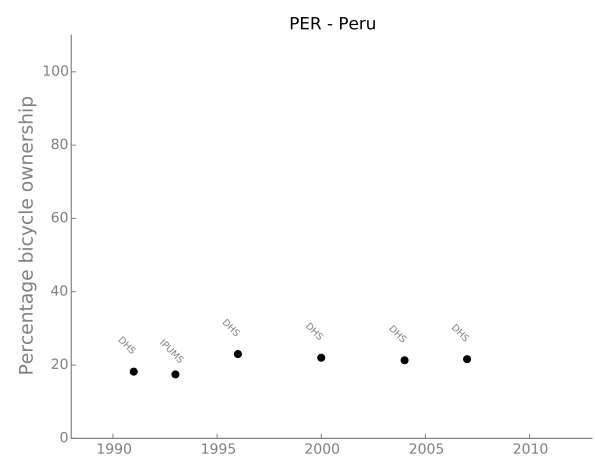
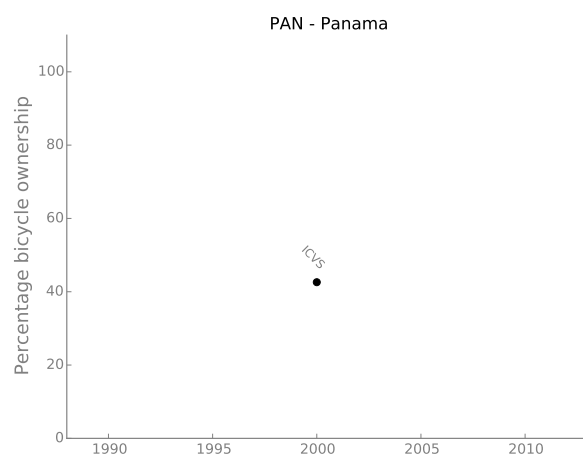
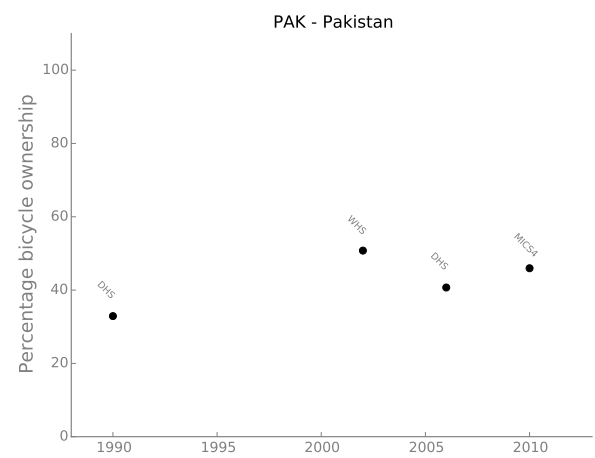
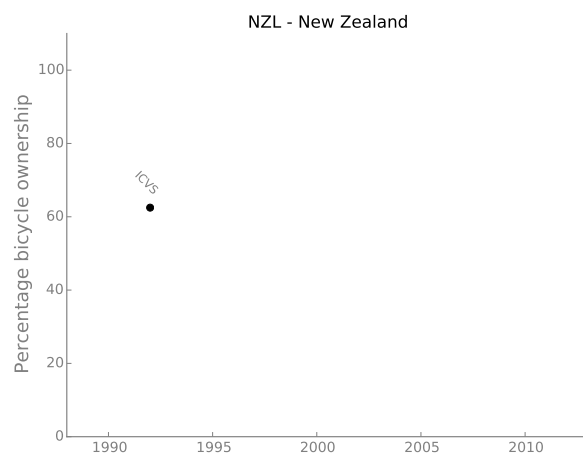
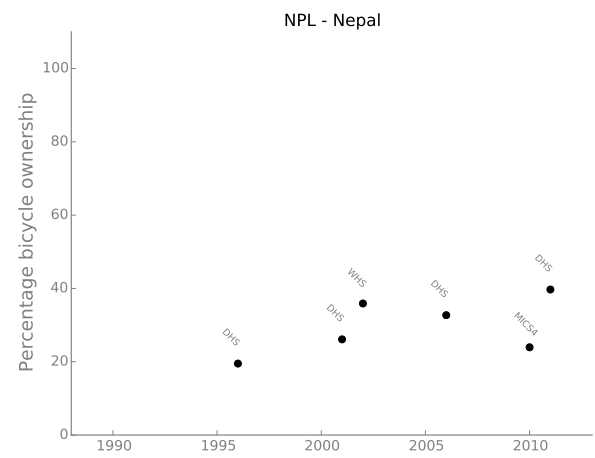
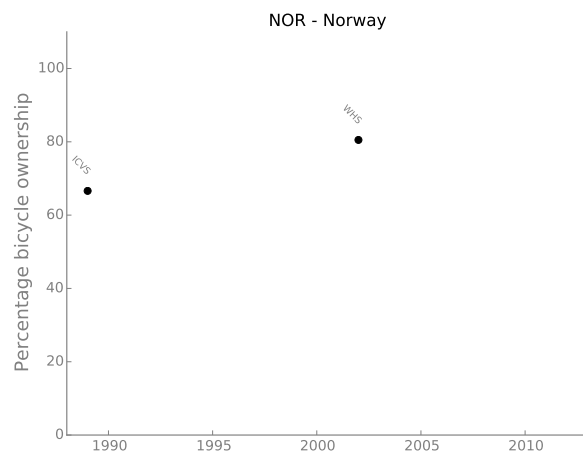
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



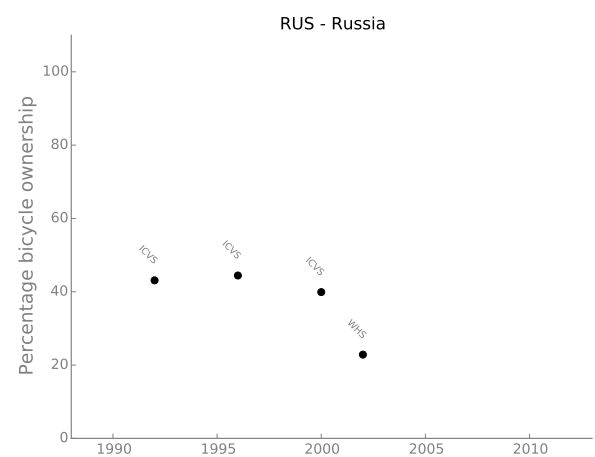
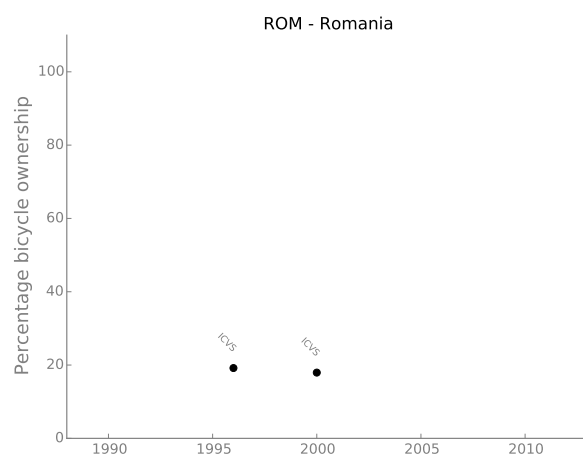
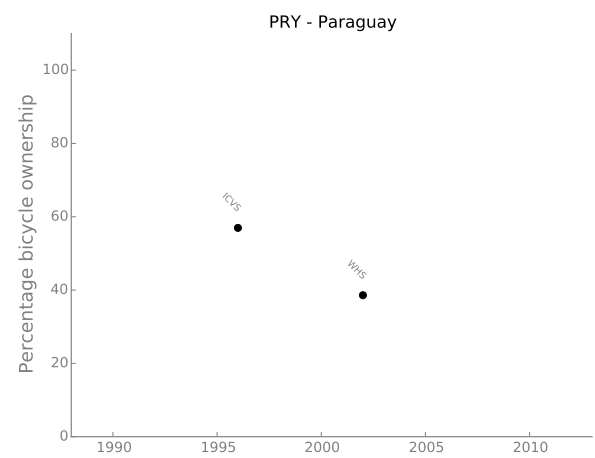
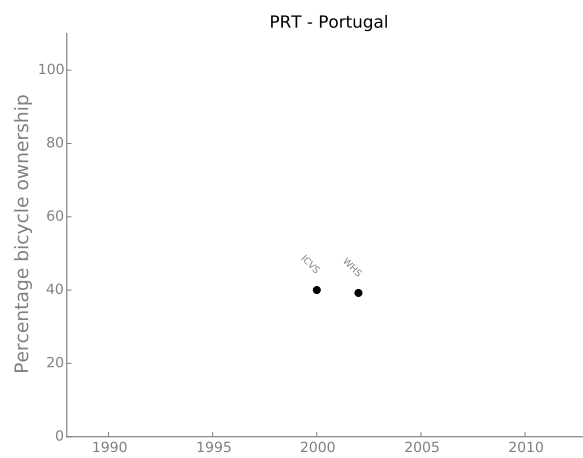
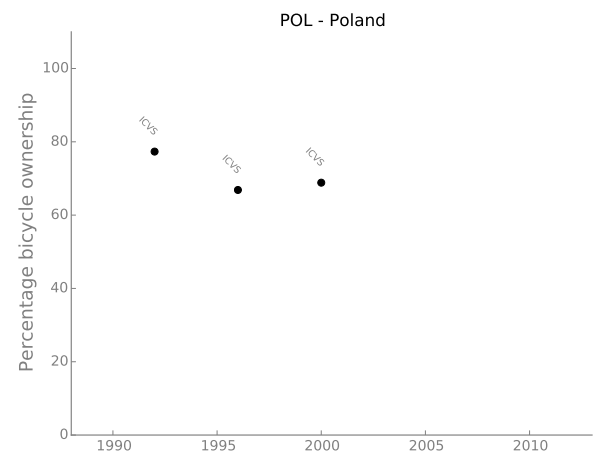
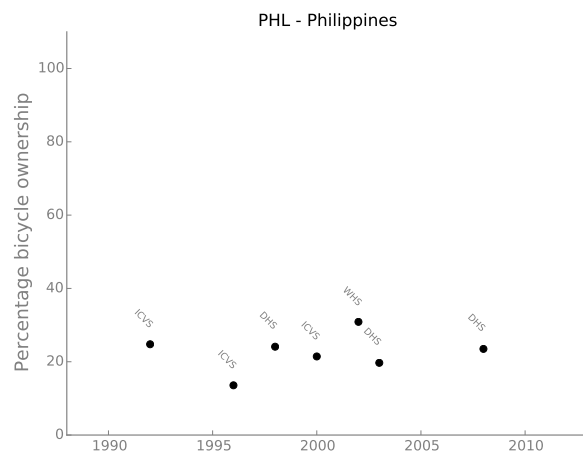
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



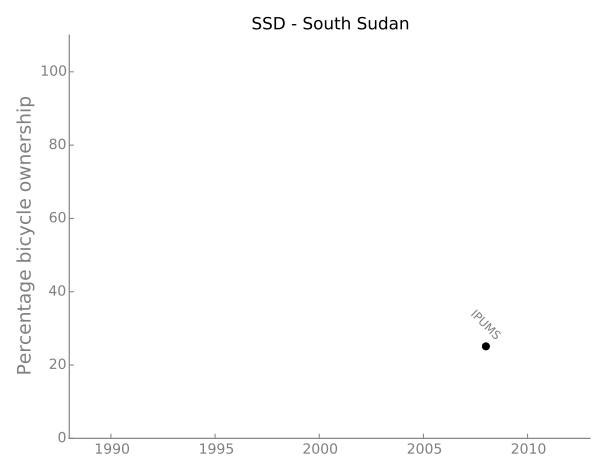
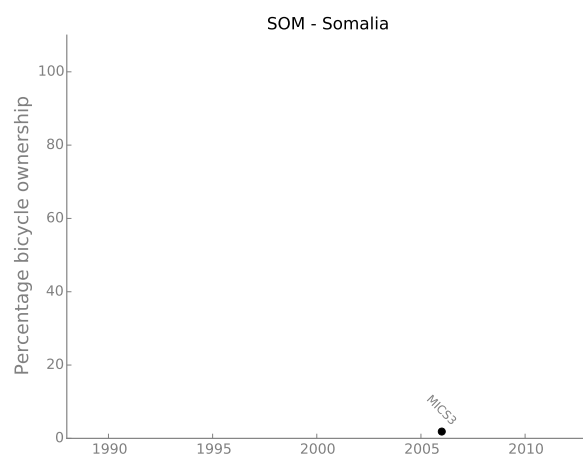
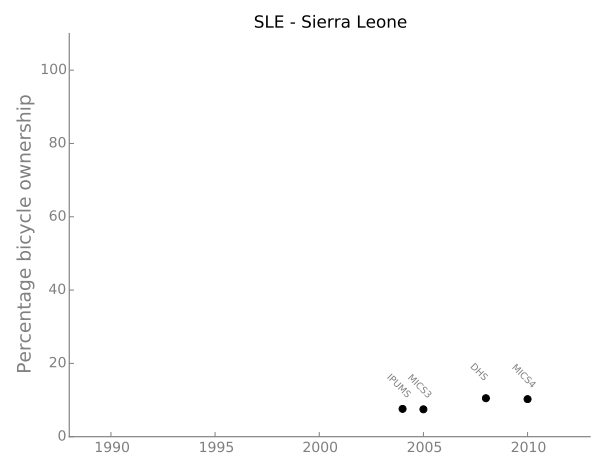
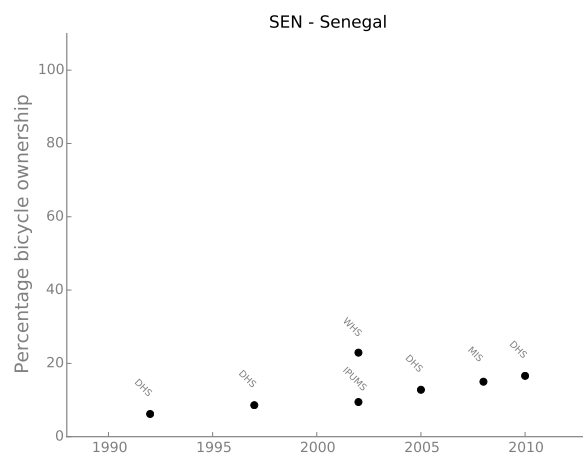
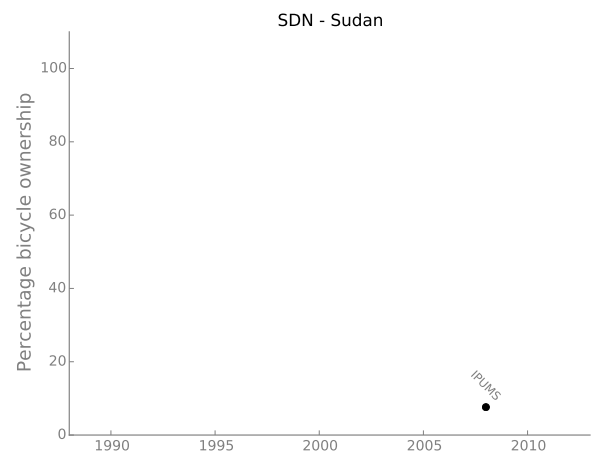
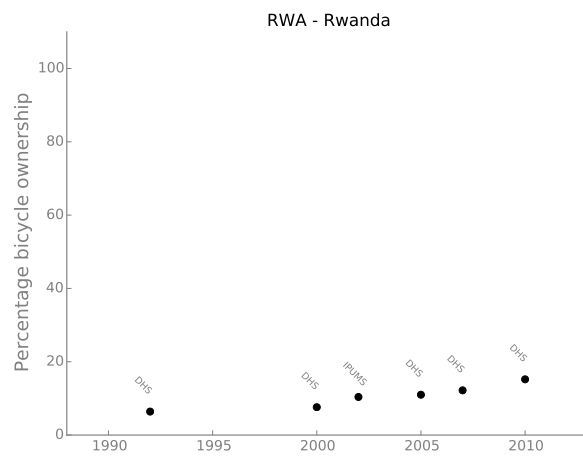
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



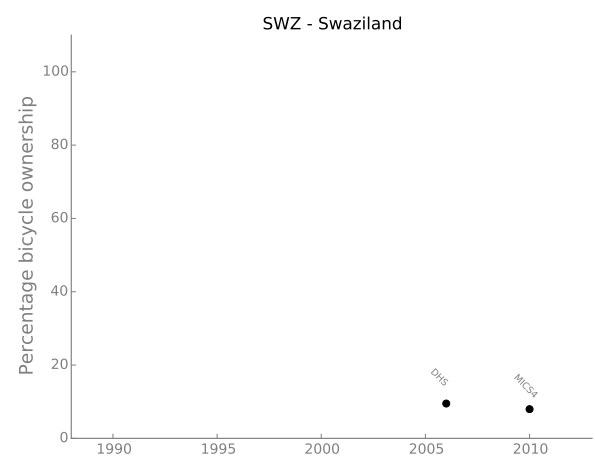
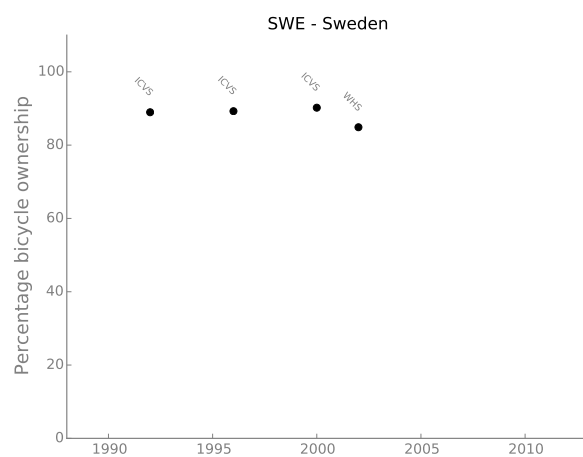
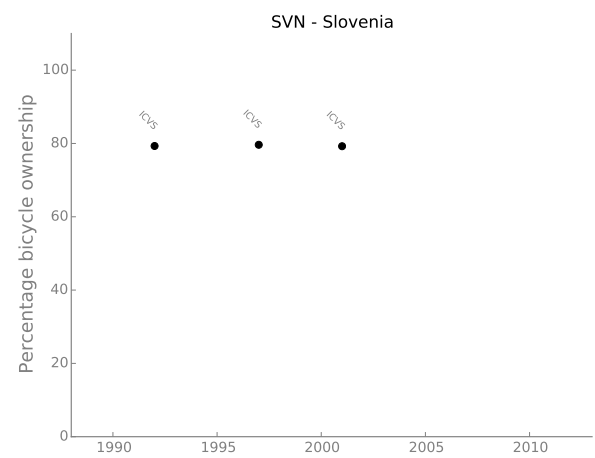
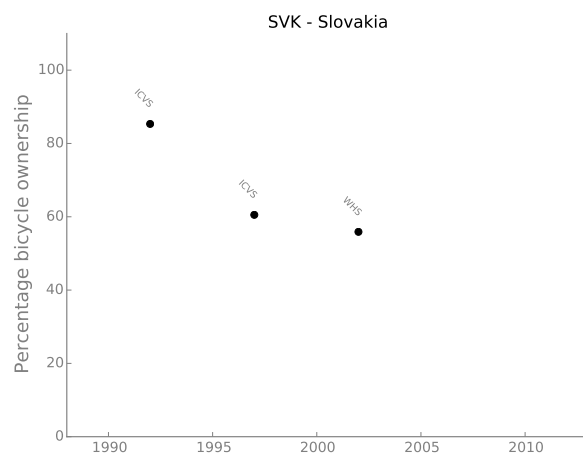
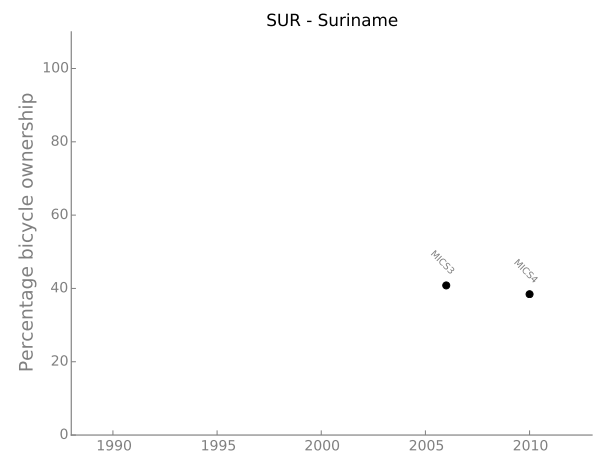
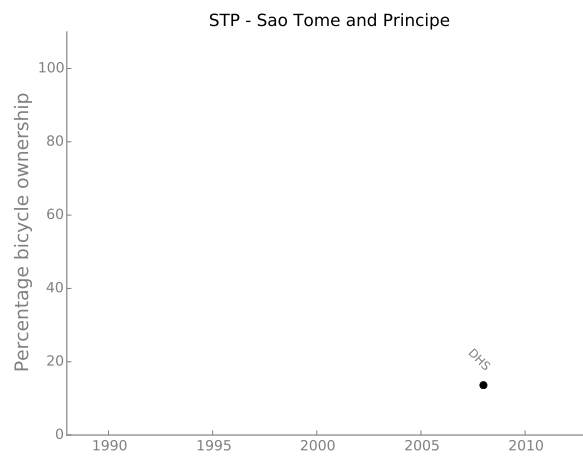
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



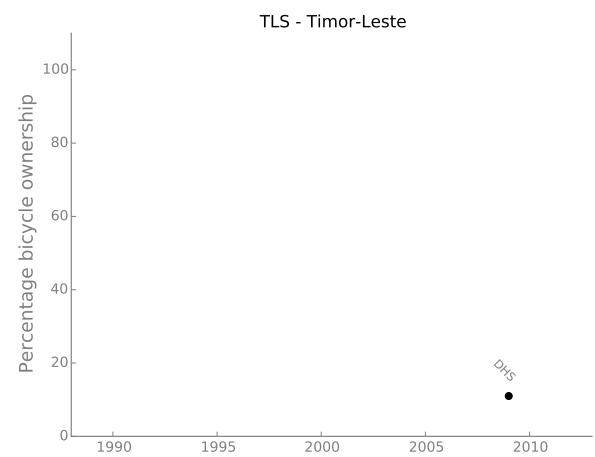
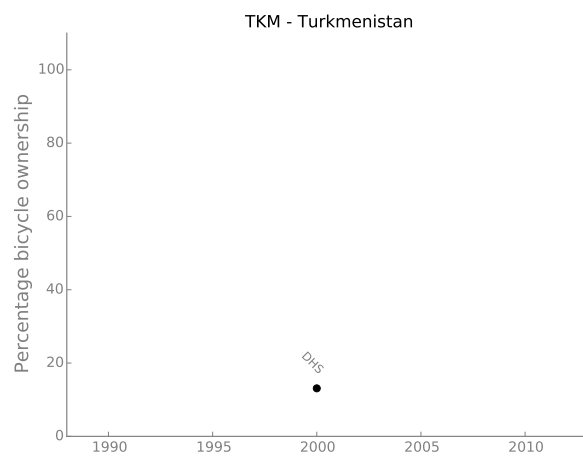
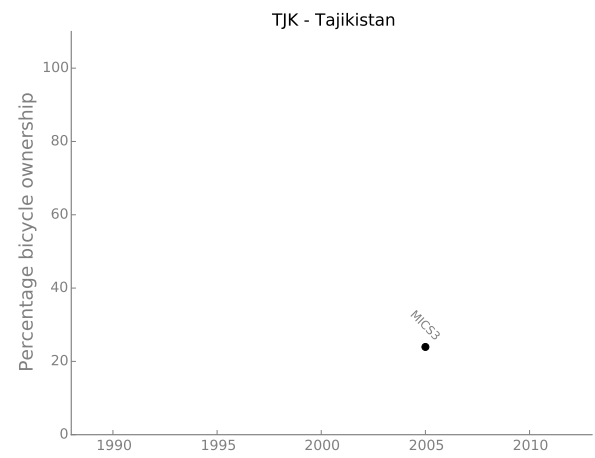
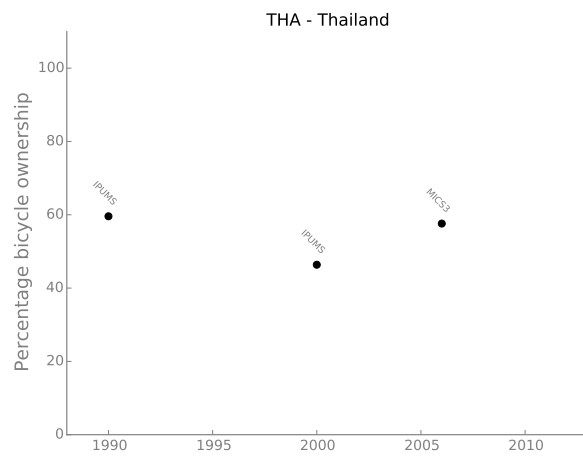
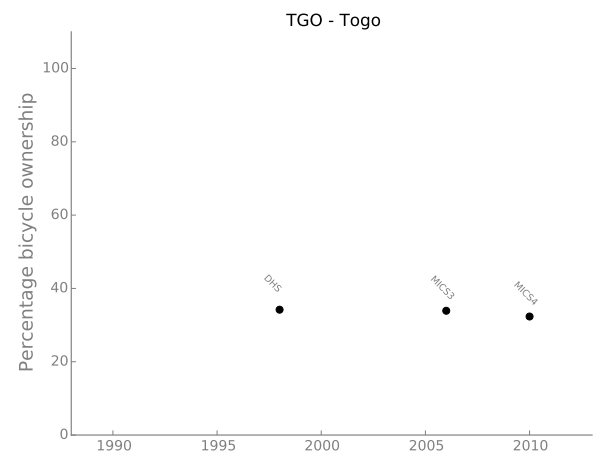
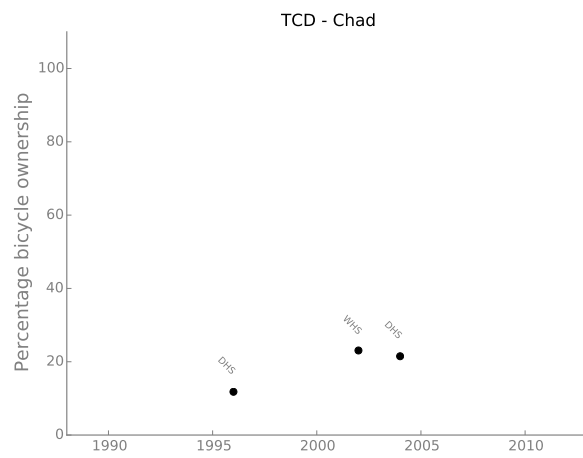
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



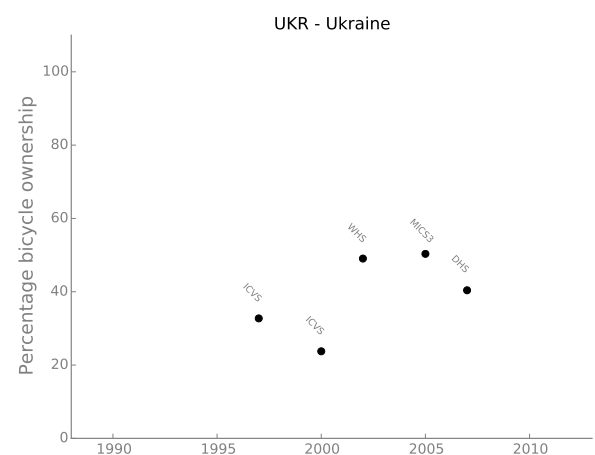
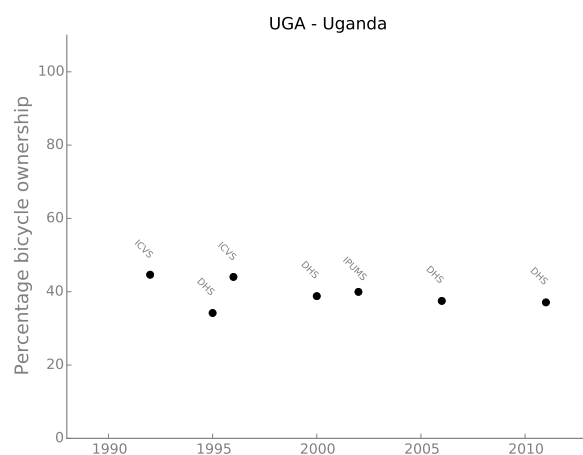
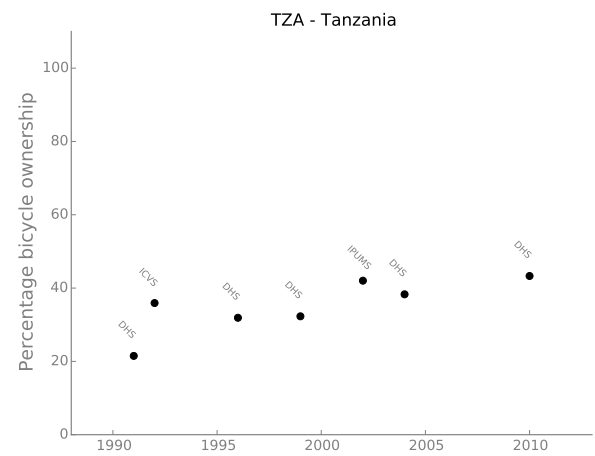
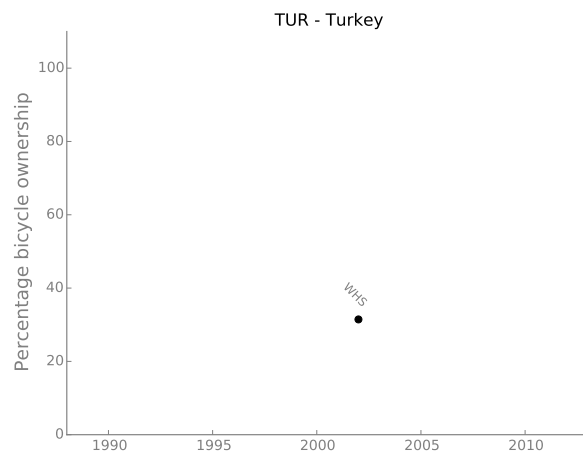
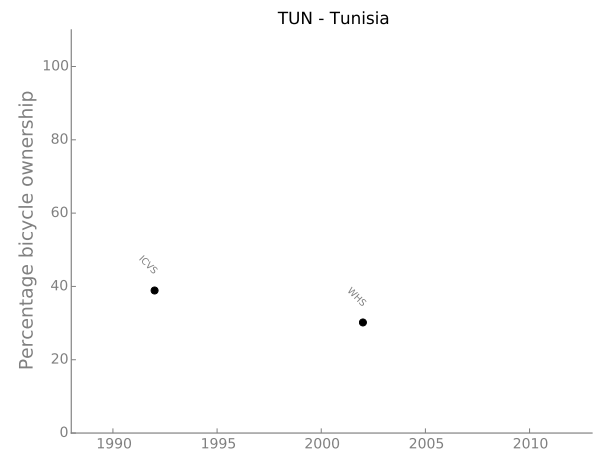
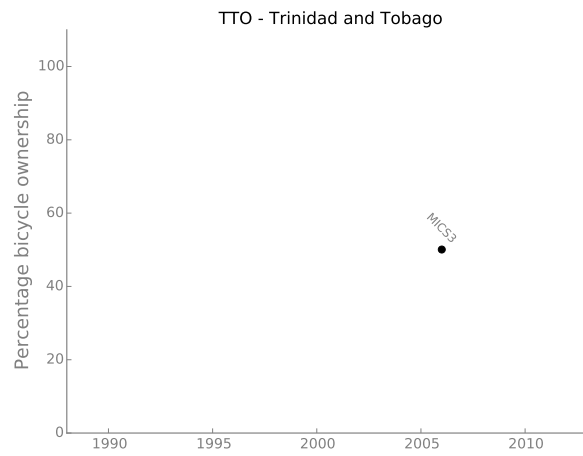
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



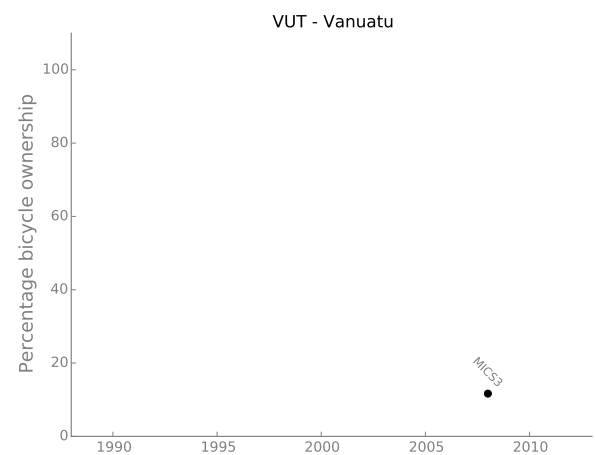
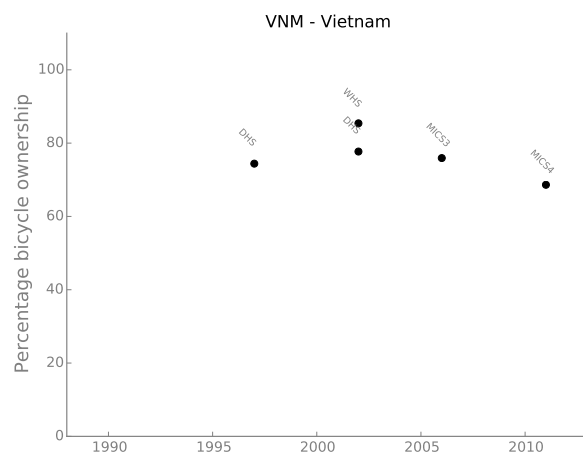
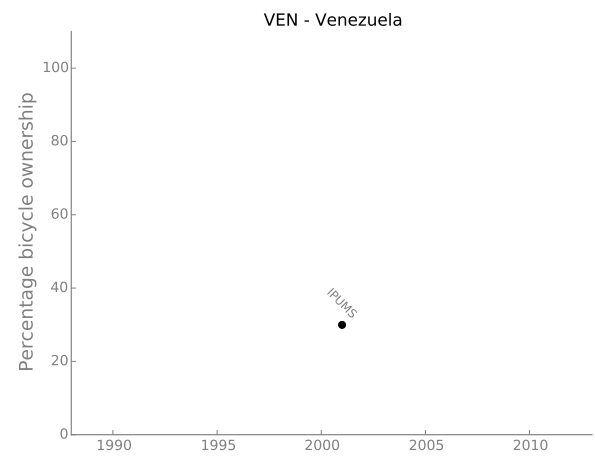
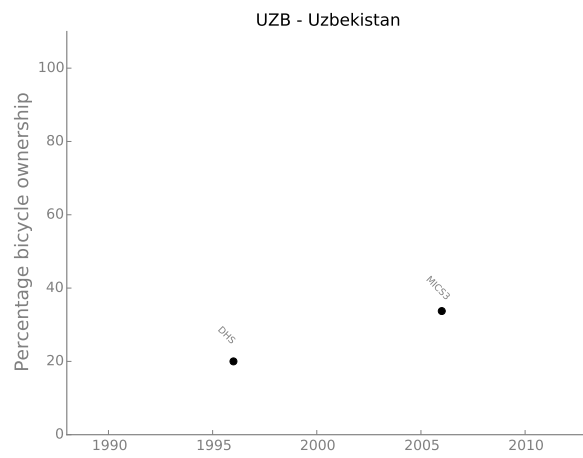
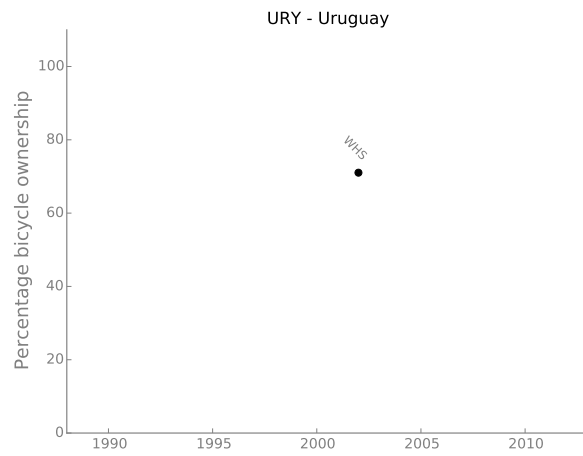
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



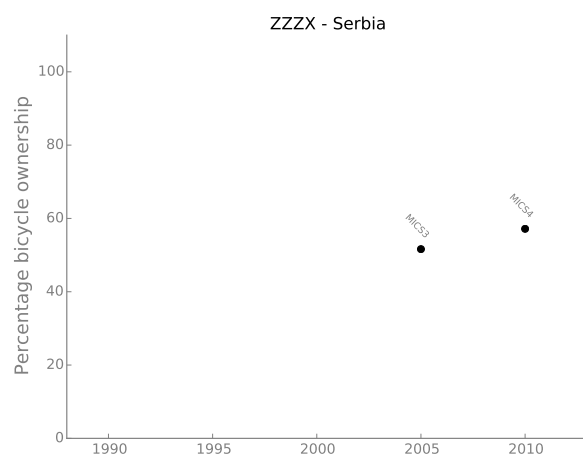
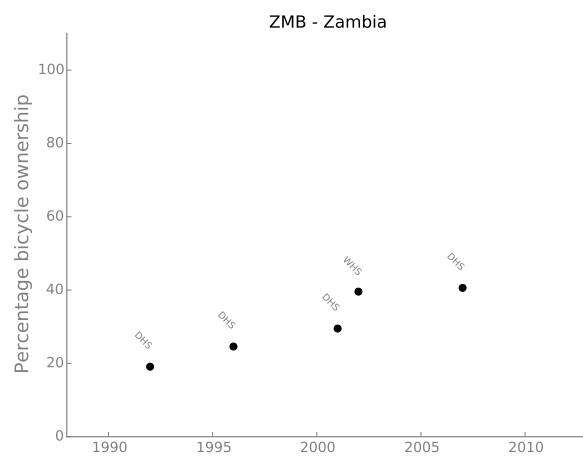
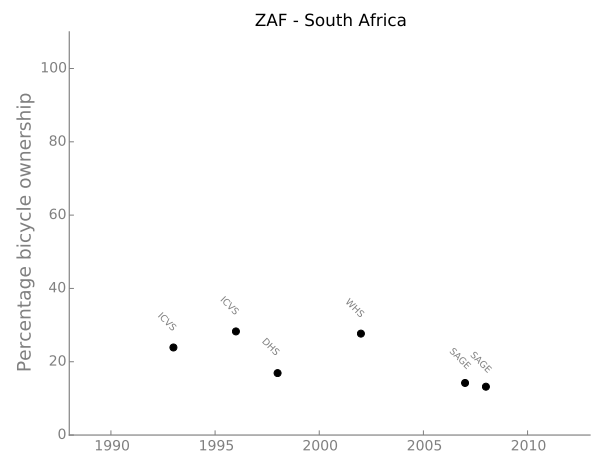
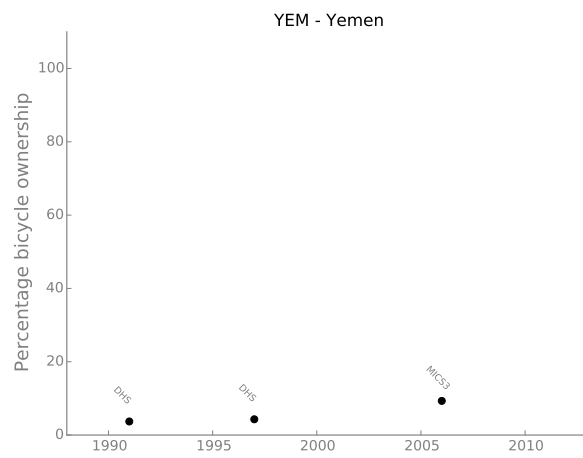
Supplementary Fig. S1 Bicycle ownership trends by country (cont)



Supplementary Fig. S1 Bicycle ownership trends by country (cont)



Supplementary Fig. S1 Bicycle ownership trends by country (cont)



Supplementary Fig. S1 Bicycle ownership trends by country (cont)

3 Reference data

3.1 Processed dataset

[Supplementary Table S3](#) is the processed reference dataset, detailing the Year, country (ISO abbreviation), percentage bicycle ownership (PBO), estimated number of households and the corresponding cluster or group number.

Supplementary Table S3 Bicycle ownership data. ‘PBO’ stands for ‘percentage bicycle ownership’

Year	ISO	PBO	Households	Group	Year	ISO	PBO	Households	Group
1989	AUS	47.00	7,760,322	2	2002	BRA	57.62	44,776,740	2
1989	BEL	54.79	3,608,178	2	2002	CHL	53.10	4,861,150	3
1989	CAN	60.72	8,281,530	2	2002	CHN	76.53	315,281,800	2
1989	CHE	62.48	2,543,540	2	2002	CIV	34.90	3,122,656	3
1989	DEU	72.27	26,977,920	1	2002	COG	8.85	575,082	4
1989	ESP	38.54	18,083,690	3	2002	COM	6.71	98,412	4
1989	FIN	82.64	1,781,771	1	2002	CZE	42.76	3,950,868	2
1989	FRA	52.31	19,392,640	2	2002	DEU	76.60	38,718,000	1
1989	GBR	31.56	20,200,000	3	2002	DNK	88.59	2,459,203	1
1989	IDN	62.52	34,315,135	2	2002	DOM	10.66	2,473,750	4
1989	JPN	70.88	51,842,307	2	2002	ECU	61.27	3,542,220	2
1989	NLD	87.86	5,111,000	1	2002	ERI	10.50	807,844	4
1989	NOR	66.61	1,523,508	1	2002	ESP	64.02	14,289,460	3
1989	USA	59.69	94,226,820	2	2002	EST	30.31	566,669	2
1990	COL	27.40	5,710,460	3	2002	ETH	2.39	15,239,290	4
1990	IDN	44.52	39,501,372	2	2002	FIN	85.00	2,354,082	1
1990	PAK	32.90	16,106,550	3	2002	FRA	60.03	23,808,072	2
1990	THA	59.59	11,813,400	2	2002	GEO	27.98	1,243,158	4
1991	CMR	15.70	1,756,460	4	2002	GHA	31.07	3,970,970	4
1991	DOM	3.80	1,222,400	4	2002	GRC	30.12	3,674,380	3
1991	IDN	40.20	39,501,372	2	2002	HRV	43.71	1,477,377	3
1991	MYS	41.34	3,909,600	3	2002	HUN	34.12	4,545,040	3
1991	PER	18.20	8,216,750	4	2002	IDN	44.20	51,249,710	2
1991	TZA	21.50	4,726,160	3	2002	IND	64.11	172,254,285	3
1991	YEM	3.70	1,845,684	4	2002	IRL	50.04	1,400,400	2
1992	ARG	67.60	11,996,510	2	2002	ISR	35.25	1,849,800	3
1992	AUS	59.07	7,760,322	2	2002	ITA	65.63	21,978,000	2
1992	BEL	65.61	3,953,125	2	2002	KAZ	28.83	4,160,216	4
1992	BRA	51.26	34,694,413	2	2002	KEN	36.33	6,342,120	4
1992	CAN	67.53	10,018,270	2	2002	LAO	53.90	1,023,107	2
1992	CHL	39.57	3,739,640	3	2002	LKA	44.69	407,065	3
1992	CHN	97.23	315,281,800	2	2002	LUX	64.04	171,953	2
1992	CRI	69.18	561,860	2	2002	LVA	17.98	802,848	3
1992	CZE	82.01	4,051,583	2	2002	MAR	31.21	4,459,460	4
1992	EGY	19.59	17,404,140	4	2002	MEX	53.21	22,268,196	3
1992	EST	65.64	597,100	2	2002	MLI	62.06	1,618,800	3
1992	FIN	83.44	2,036,732	1	2002	MMR	52.84	10,616,662	2
1992	GBR	50.39	21,576,100	3	2002	MOZ	22.15	4,655,396	4
1992	GEO	18.93	1,243,158	4	2002	MRT	4.92	368,429	4
1992	IDN	56.59	39,501,372	2	2002	MUS	75.34	344,785	2
1992	IND	25.56	143,400,000	3	2002	MWI	39.16	2,290,050	3
1992	ITA	63.53	25,007,000	2	2002	MYS	44.49	4,778,200	3
1992	MAR	9.90	3,433,680	4	2002	NAM	18.43	16,839	4
1992	MDG	3.50	2,666,065	4	2002	NLD	89.48	6,934,263	1
1992	MWI	21.20	1,862,700	3	2002	NOR	80.51	1,961,548	1
1992	NAM	19.70	21,283	4	2002	NPL	35.89	4,586,995	4
1992	NER	4.70	1,134,000	4	2002	PAK	50.77	19,579,910	3
1992	NLD	90.35	6,061,000	1	2002	PHL	30.88	15,118,900	4
1992	NZL	62.49	1,443,636	2	2002	PRT	39.22	5,176,860	3
1992	PHL	24.77	11,561,260	4	2002	PRY	38.60	1,238,082	3
1992	POL	77.33	11,970,440	2	2002	RUS	22.85	52,711,375	3
1992	RUS	43.11	54,560,627	3	2002	RWA	10.39	1,917,190	4
1992	RWA	6.40	1,530,410	4	2002	SEN	16.20	1,079,990	4
1992	SEN	6.20	799,040	4	2002	SVK	55.89	2,071,743	2
1992	SVK	85.33	1,832,484	2	2002	SWE	84.88	4,448,746	1
1992	SVN	79.31	640,195	1	2002	TCD	23.07	1,118,597	4
1992	SWE	88.97	4,141,450	1	2002	TUN	30.17	1,458,100	3
1992	TUN	38.89	1,458,100	3	2002	TUR	31.45	16,446,644	3
1992	TZA	35.92	4,726,160	3	2002	TZA	42.02	8,417,680	3
1992	UGA	44.64	3,391,660	3	2002	UGA	39.96	5,292,710	3
1992	ZMB	19.10	1,295,010	3	2002	UKR	49.05	17,609,200	3
1993	BFA	66.80	1,452,690	1	2002	URY	71.03	1,180,670	2
1993	BGD	15.90	19,711,240	4	2002	VNM	81.55	17,804,633	2
1993	ESP	29.85	11,845,520	3	2002	ZAF	27.67	11,202,529	4
1993	GHA	16.10	5,702,340	4	2002	ZMB	39.58	1,862,250	3
1993	KEN	22.10	4,497,220	4	2002	ZWE	34.28	2,834,287	4
1993	PER	17.45	5,647,650	4	2003	BFA	78.00	1,631,770	1
1993	ZAF	23.89	17,258,500	4	2003	BOL	43.60	2,394,750	3
1994	CAF	11.50	802,163	4	2003	GHA	23.10	3,970,970	4
1994	CIV	26.20	2,584,056	3	2003	HTI	10.49	2,196,330	4
1994	ESP	24.48	11,845,520	3	2003	KEN	29.30	6,342,120	4
1994	HTI	9.20	1,147,920	4	2003	MAR	25.40	4,459,460	4
1994	IDN	45.40	39,501,372	2	2003	MDG	16.00	3,748,972	4
1994	ZWE	18.00	2,562,478	4	2003	MOZ	32.60	4,788,785	4

Supplementary Table S3 Bicycle ownership data. ‘PBO’ stands for ‘percentage bicycle ownership’

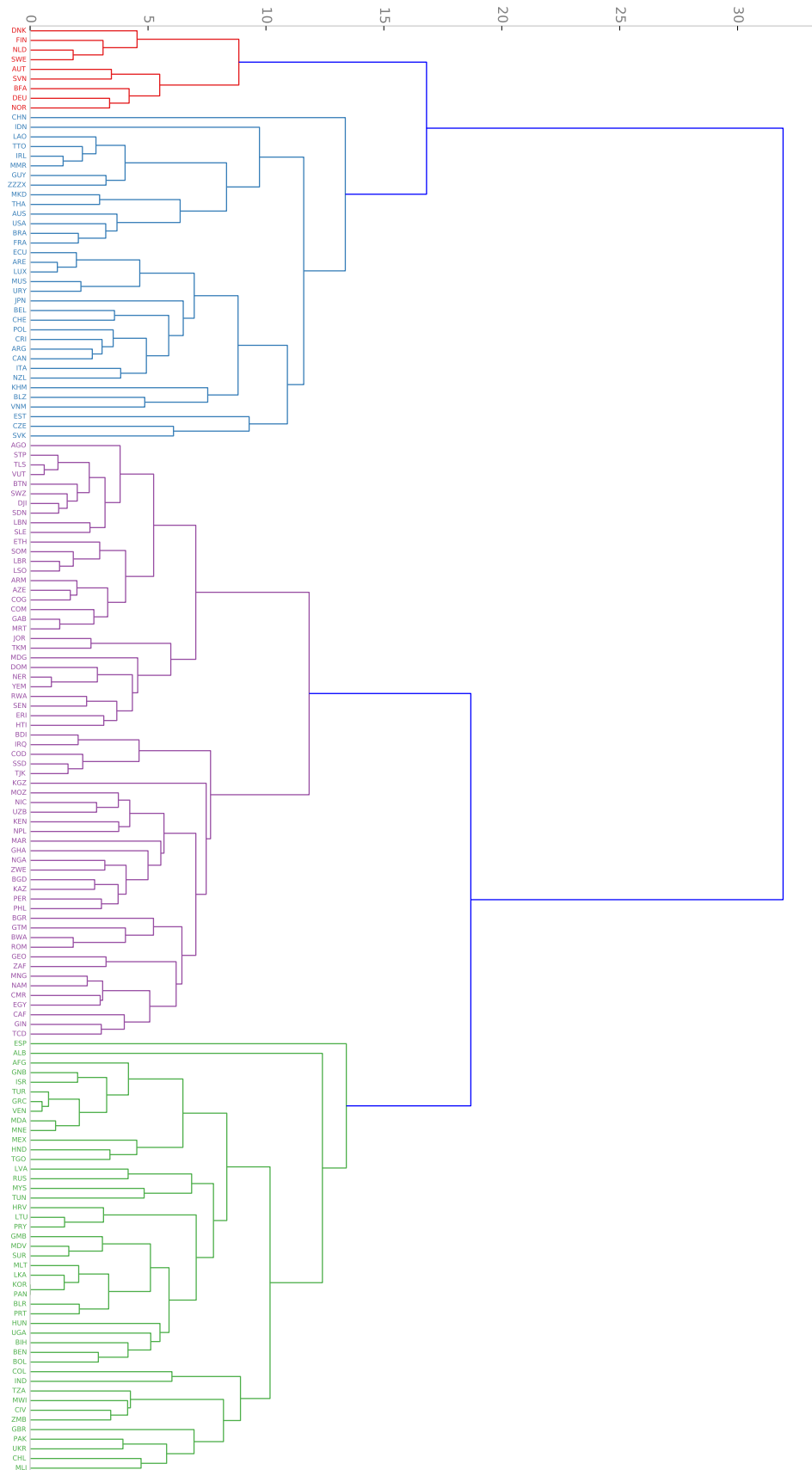
Year	ISO	PBO	Households	Group	Year	ISO	PBO	Households	Group
1995	COL	41.50	7,743,210	3	2003	NGA	32.70	43,404,000	4
1995	EGY	14.70	17,404,140	4	2003	PHL	19.70	15,118,900	4
1995	ERI	7.20	642,983	4	2004	BGD	24.20	26,259,590	4
1995	EST	67.27	597,100	2	2004	CMR	18.20	1,756,460	4
1995	GTM	22.80	1,980,924	4	2004	LSO	3.00	422,371	4
1995	IDN	52.98	38,612,325	2	2004	MWI	40.00	2,290,050	3
1995	KAZ	17.30	4,391,759	4	2004	PER	21.30	5,647,650	4
1995	MLI	37.20	1,365,150	3	2004	SLE	7.59	825,180	4
1995	UGA	34.20	3,391,660	3	2004	TCD	21.50	1,206,619	4
1996	ALB	50.22	791,830	3	2004	TZA	38.30	8,417,680	3
1996	ARG	67.41	11,996,510	2	2005	ALB	22.30	733,860	3
1996	AUT	81.14	3,093,200	1	2005	ARM	5.20	819,290	4
1996	BEN	42.60	764,000	3	2005	BDI	13.01	124,911	4
1996	BGD	19.30	19,711,240	4	2005	COD	24.20	7,945,294	4
1996	BOL	49.61	1,779,260	3	2005	COG	5.90	620,467	4
1996	BRA	61.34	34,694,413	2	2005	EGY	14.50	12,707,870	4
1996	CAN	66.62	10,820,055	2	2005	ETH	1.20	16,594,169	4
1996	CHE	76.21	2,958,940	2	2005	GEO	8.16	1,173,675	4
1996	COM	2.50	84,521	4	2005	GIN	25.00	1,107,770	4
1996	CRI	65.06	561,860	2	2005	HND	39.30	1,444,000	3
1996	CZE	70.30	3,692,134	2	2005	HTI	17.50	2,196,330	4
1996	DOM	3.30	1,222,400	4	2005	IDN	44.28	52,300,392	2
1996	FIN	91.00	2,180,934	1	2005	IND	51.10	207,800,000	3
1996	FRA	61.97	22,616,500	2	2005	KHM	68.30	2,235,130	2
1996	GBR	49.30	21,576,100	3	2005	MDA	27.20	1,391,655	3
1996	GEO	28.31	1,243,158	4	2005	MKD	57.66	501,963	2
1996	HUN	45.90	4,387,780	3	2005	MNE	29.36	180,517	3
1996	IDN	68.80	38,612,325	2	2005	MNG	15.77	557,950	4
1996	IND	36.83	164,870,000	3	2005	NIC	31.99	1,193,390	4
1996	KGZ	35.08	2,493,770	4	2005	RWA	11.00	1,917,190	4
1996	LVA	36.41	859,823	3	2005	SEN	12.80	1,079,990	4
1996	MKD	63.11	501,963	2	2005	SLE	7.47	825,180	4
1996	MKD	48.11	501,963	2	2005	TJK	23.94	1,047,000	4
1996	MNG	18.05	427,830	4	2005	UKR	50.34	17,417,500	3
1996	NLD	91.13	6,468,682	1	2005	ZWE	25.10	2,849,907	4
1996	NPL	19.50	4,586,995	4	2005	ZZZX	51.62	2,521,190	2
1996	PER	23.00	5,647,650	4	2006	AGO	17.90	3,567,168	4
1996	PHL	13.57	13,621,900	4	2006	AZE	6.90	1,817,460	4
1996	POL	66.86	12,501,802	2	2006	BEN	39.50	936,000	3
1996	PRY	56.97	1,091,489	3	2006	BFA	80.72	2,362,060	1
1996	ROM	19.18	7,288,460	4	2006	BGD	24.98	28,716,961	4
1996	RUS	44.43	54,560,627	3	2006	BIH	36.10	1,067,120	3
1996	SWE	89.26	4,243,980	1	2006	BLZ	71.55	79,272	2
1996	TCD	11.80	900,913	4	2006	CAF	18.89	1,010,551	4
1996	TZA	31.90	4,726,160	3	2006	CIV	43.14	3,307,568	3
1996	UGA	44.04	3,391,660	3	2006	CMR	15.23	3,453,630	4
1996	USA	59.68	98,990,000	2	2006	DJI	8.98	125,006	4
1996	UZB	20.00	4,466,346	4	2006	EGY	16.28	17,404,140	4
1996	ZAF	28.27	9,938,010	4	2006	GHA	3.47	3,970,970	4
1996	ZMB	24.60	1,295,010	3	2006	GMB	44.24	158,032	3
1996	ZWE	27.11	2,656,078	4	2006	GNB	35.73	274,086	3
1997	BGR	32.50	2,964,577	4	2006	GUY	55.17	188,979	2
1997	BLR	36.40	3,873,139	3	2006	IRQ	15.01	2,654,020	4
1997	BWA	20.42	404,706	4	2006	KAZ	21.78	4,160,216	4
1997	COL	72.83	7,743,210	3	2006	KGZ	23.19	1,109,633	4
1997	HRV	56.72	1,544,250	3	2006	LAO	47.21	1,087,809	2
1997	IDN	47.50	38,612,325	2	2006	LBN	3.76	1,019,955	4
1997	JOR	8.90	973,430	4	2006	MLI	46.20	1,618,800	3
1997	KGZ	10.30	2,493,770	4	2006	MWI	43.21	2,290,050	3
1997	LTU	55.49	1,356,826	3	2006	NAM	15.20	16,839	4
1997	MDG	5.30	3,115,051	4	2006	NER	10.20	1,559,000	4
1997	MLT	41.47	119,479	3	2006	NPL	32.70	4,586,995	4
1997	MOZ	15.00	4,075,813	4	2006	PAK	40.70	19,579,910	3
1997	SEN	8.60	799,040	4	2006	SOM	1.85	1,860,314	4
1997	SVK	60.53	1,832,484	2	2006	SUR	40.82	134,001	3
1997	SVN	79.64	640,195	1	2006	SWZ	9.50	172,414	4
1997	UKR	32.73	36,235,710	3	2006	TGO	33.91	1,184,551	3
1997	VNM	74.40	10,684,460	2	2006	THA	57.60	16,541,700	2
1997	YEM	4.30	2,397,618	4	2006	TTO	50.07	352,291	2
1998	BFA	77.20	1,631,770	1	2006	UGA	37.50	5,292,710	3
1998	CIV	29.60	2,890,724	3	2006	UZB	33.73	5,093,894	4
1998	CMR	13.30	1,756,460	4	2006	VNM	75.92	17,804,633	2
1998	GHA	17.40	5,702,340	4	2006	YEM	9.32	3,079,242	4
1998	GTM	26.70	2,121,248	4	2007	BGD	26.30	28,716,961	4
1998	IND	47.80	164,870,000	3	2007	CHN	48.65	2,226,546	2
1998	KEN	23.90	4,497,220	4	2007	DOM	1.50	2,473,750	4
1998	NER	4.80	1,246,000	4	2007	GHA	24.10	3,970,970	4
1998	NGA	6.14	43,404,000	4	2007	IDN	47.20	52,300,392	2
1998	NIC	18.50	828,150	4	2007	IND	64.38	207,800,000	3
1998	PHL	24.10	13,621,900	4	2007	LBR	4.10	688,360	4
1998	TGO	34.20	962,647	3	2007	NGA	27.62	36,840,000	4
1998	ZAF	16.90	9,938,010	4	2007	PER	21.60	8,216,750	4
1999	BGD	20.30	19,711,240	4	2007	RWA	12.20	1,917,190	4
1999	DOM	4.10	1,222,400	4	2007	UKR	40.40	17,341,100	3
1999	GIN	14.90	1,107,770	4	2007	ZAF	14.21	17,258,500	4
1999	KAZ	14.00	4,391,759	4	2007	ZMB	40.60	1,862,250	3

Supplementary Table S3 Bicycle ownership data. ‘PBO’ stands for ‘percentage bicycle ownership’

Year	ISO	PBO	Households	Group	Year	ISO	PBO	Households	Group
1999	NGA	24.20	43,404,000	4	2008	AGO	8.70	3,815,508	4
1999	TZA	32.30	4,726,160	3	2008	ALB	21.40	791,830	3
1999	ZWE	20.00	2,776,844	4	2008	BOL	36.60	2,394,750	3
2000	ALB	57.03	791,830	3	2008	CHN	60.87	2,226,546	2
2000	ARM	6.60	819,290	4	2008	EGY	9.80	17,404,140	4
2000	AUS	57.99	7,760,322	2	2008	GHA	33.86	3,970,970	4
2000	AZE	5.40	1,706,909	4	2008	IND	33.80	207,800,000	3
2000	BEL	72.90	4,237,775	2	2008	KEN	30.10	6,342,120	4
2000	BGR	22.62	2,964,577	4	2008	KHM	67.99	2,895,620	2
2000	BLR	38.49	3,855,016	3	2008	MDG	22.60	4,331,909	4
2000	CAN	67.28	10,820,055	2	2008	MOZ	40.05	5,484,945	4
2000	CHE	73.43	3,303,540	2	2008	MWI	44.83	2,998,640	3
2000	COL	60.51	7,743,210	3	2008	NGA	22.90	43,404,000	4
2000	CZE	58.43	3,854,791	2	2008	PHL	23.50	15,118,900	4
2000	DNK	87.52	2,438,307	1	2008	SDN	7.61	5,972,988	4
2000	EGY	14.00	12,707,870	4	2008	SEN	15.00	1,079,990	4
2000	EST	66.33	582,089	2	2008	SLE	10.50	825,180	4
2000	ETH	0.80	14,384,357	4	2008	SSD	25.10	1,322,742	4
2000	FIN	94.69	2,295,386	1	2008	STP	13.60	33,772	4
2000	FRA	64.44	23,808,072	2	2008	VUT	11.66	42,527	4
2000	GAB	3.40	194,000	4	2008	ZAF	13.19	1,322,742	4
2000	GBR	53.55	24,396,000	3	2009	CHN	63.12	2,226,546	2
2000	GEO	11.74	1,107,722	4	2009	GUY	52.10	192,853	2
2000	HRV	58.27	1,544,250	3	2009	KEN	29.45	8,952,300	4
2000	HTI	13.90	1,147,920	4	2009	LBR	3.00	688,360	4
2000	HUN	52.39	4,387,780	3	2009	LSO	2.10	422,371	4
2000	KHM	52.50	2,235,130	2	2009	MDV	39.90	43,197	3
2000	KOR	42.61	17,339,422	3	2009	MEX	36.93	25,469,850	3
2000	LTU	38.94	1,356,826	3	2009	TLS	11.00	194,962	4
2000	LVA	42.85	802,848	3	2009	ZWE	21.73	2,889,891	4
2000	MNG	21.10	557,950	4	2010	ARM	4.00	819,290	4
2000	MWI	43.40	2,290,050	3	2010	BDI	20.50	124,911	4
2000	MYS	33.17	4,778,200	3	2010	BFA	84.20	2,362,060	1
2000	NAM	17.60	21,283	4	2010	BTN	6.18	127,942	4
2000	NLD	90.90	6,801,008	1	2010	CHN	62.97	2,226,546	2
2000	PAN	42.60	843,460	3	2010	COD	21.26	9,145,758	4
2000	PER	22.00	5,647,650	4	2010	COL	33.30	10,570,899	3
2000	PHL	21.44	15,118,900	4	2010	KHM	65.50	2,895,620	2
2000	POL	68.85	13,431,700	2	2010	MEX	32.88	29,036,410	3
2000	PRT	40.01	4,283,100	3	2010	MWI	43.80	2,998,640	3
2000	ROM	17.94	7,288,460	4	2010	NPL	23.94	4,586,995	4
2000	RUS	39.90	54,560,627	3	2010	PAK	45.96	19,579,910	3
2000	RWA	7.60	1,530,410	4	2010	RWA	15.20	1,917,190	4
2000	SWE	90.21	4,363,284	1	2010	SEN	16.60	1,079,990	4
2000	THA	46.37	16,541,700	2	2010	SLE	10.25	825,180	4
2000	TKM	13.10	882,631	4	2010	SUR	38.42	139,246	3
2000	UGA	38.80	3,391,660	3	2010	SWZ	7.96	172,414	4
2000	UKR	23.75	17,679,600	3	2010	TGO	32.34	1,313,752	3
2000	USA	56.89	123,688,760	2	2010	TZA	43.30	8,417,680	3
2001	ARG	72.49	10,408,520	2	2010	ZWE	22.60	2,932,057	4
2001	BEN	45.50	841,000	3	2010	ZZZX	57.16	2,521,190	2
2001	BOL	43.78	2,394,750	3	2011	AFG	31.82	3,933,172	3
2001	IND	43.67	172,254,285	3	2011	BGD	25.40	33,092,620	4
2001	KHM	62.95	2,235,130	2	2011	BLZ	70.92	79,272	2
2001	MLI	49.50	1,618,800	3	2011	CMR	14.70	3,453,630	4
2001	NIC	25.70	828,150	4	2011	ETH	2.30	19,475,612	4
2001	NPL	26.10	4,586,995	4	2011	HND	33.30	1,707,000	3
2001	SVN	79.25	640,195	1	2011	IND	44.82	207,800,000	3
2001	VEN	29.96	6,460,800	3	2011	IRQ	23.28	2,654,020	4
2001	ZMB	29.50	1,862,250	3	2011	KAZ	24.78	4,391,759	4
2002	ARE	66.35	642,224	2	2011	LBN	10.10	1,095,697	4
2002	AUT	84.85	3,314,200	1	2011	NGA	21.57	43,404,000	4
2002	BEL	72.77	4,319,040	2	2011	NPL	39.70	4,586,995	4
2002	BFA	79.07	1,631,770	1	2011	UGA	37.10	5,292,710	3
2002	BGD	26.71	26,259,590	4	2011	VNM	68.63	24,613,613	2
2002	BIH	47.61	1,206,791	3	2012	BIH	45.58	1,206,791	3

3.2 Unweighted pair-group method with arithmetic mean: dendrogram

As discussed in the paper, the unweighted pair-group method was the best-fit clustering method for our data. The dendrogram produced by the method is shown in [Supplementary Fig. S2](#).



Supplementary Fig. S2 Dendrogram of UPGMA clustering

4 Code organization

The code for analyzing these data was written primarily in Python (using the IPython [17] system, along with the Pandas [18] and Matplotlib [19] packages), and it is available with other data resources at ce.jhu.edu/sauleh/obls-gbu.

Along with the main file (`main-v3.py`), supporting modules contain functions to carry out the tasks as listed below:

- (1) `initialize.py`: Initialize data into dataframe
- (2) `alignment.py`: Perform alignment using DTW algorithm and obtain separation matrix
- (3) `clustering.py`: Hosts the functions that perform the goodness-of-fit and gap tests, and then perform the actual clustering to put the countries in groups
- (4) `household.py`: Interpolation and data-reading functions to compute household statistics for each country
- (5) `viz.py`: Visualization functions (maps, plots, etc) to make sense of data and results.
- (6) `summaries.py`

An R script—`gapTest.R`—was also written to execute the gap test, but it is called from within the `gapTest` function in the `clustering` module.

The raw bicycle ownership data (organized into a spreadsheet) and other required elements (household data, shape files, etc) are stored in the `resources` directory. All tabular output is kept in `tables`, and visual output is stored in `Images`.

References

- [1] WHS. World Health Survey. <http://apps.who.int/healthinfo/systems/surveydata/index.php/catalog/whs>, June 2013. World Health Organization; accessed on May 27, 2013.
- [2] DHS. The Demographic and Health Surveys Program. <http://www.statcompiler.com>, June 2014. Accessed on May 27, 2013.
- [3] MIS. Malaria Indicator Surveys. <http://www.dhsprogram.com/What-We-Do/Survey-Types/MIS.cfm>, June 2014. The Demographic and Health Surveys Program. Accessed on May 27, 2013.
- [4] IPUMS. Integrated Public Use Microdata Series, International: Version 6.2 [Machine Readable Database]. <https://international.ipums.org/international/>, 2013. accessed on June 13, 2013.
- [5] ICVS. International Crime Victimization Survey (ICVS) 1989-2000. http://www.unicri.it/services/library_documentation/publications/icvs/data/, 2000. ICVS International Working Group, Anna Alvazzi del Frate, Jan J.M. van Dijk, John van Kesteren, Pat Mayhew, Ugi Svekic, United Nations Interregional Crime and Justice Research Institute (UNICRI), Turin, Italy and United Nations Office of Drugs and Crime, Vienna, Austria.
- [6] IndiaStat. India National Census. <http://www.indiastat.com>, June 2014. accessed on June 10, 2013.
- [7] MICS4. Multiple Indicator Cluster Surveys/MICS4. http://www.childinfo.org/mics4_surveys.html, April 2013. UNICEF Child Info. Accessed on June 14, 2013.
- [8] MICS3. Multiple Indicator Cluster Surveys. http://www.childinfo.org/mics3_surveys.html, April 2013. UNICEF Child Info. Accessed on June 14, 2013.
- [9] SAGE. Study on global ageing and adult health. <http://apps.who.int/healthinfo/systems/surveydata/index.php/catalog/sage>, February 2012. World Health Organization. Accessed on July 1, 2013.
- [10] IBEP. Inquérito Integrado Sobre o Bem-Estar da População. http://www.childinfo.org/mics3_surveys.html, April 2013. UNICEF Child Info. Accessed on June 14, 2013.
- [11] EDSM. Enquête Démographique et de Santé et à Indicateurs Multiples. http://www.childinfo.org/mics3_surveys.html, April 2013. The Demographic and Health Surveys Program. Accessed on June 14, 2013.
- [12] SACMEQ. Data archive for the SACMEQ III project (v.6.0). Compact Disc, 2012. Southern and Eastern Africa Consortium for Monitoring Educational Quality. Compiled by: Demus Makuwa, Njora Hungi, Kenneth N. Ross, Mioko Saito, Stephanie Dolata, Frank van Capelle, Laura Paviot, Jocelyne Vellien, Jan Maarse.
- [13] SACMEQ. Data archive for the SACMEQ I and SACMEQ II projects (v.6.0). Compact Disc, October 2009. Southern and Eastern Africa Consortium for Monitoring Educational Quality. Compiled by: Kenneth N. Ross, Mioko Saito, Stephanie Dolata, Miyako Ikeda and Linda Zuze.
- [14] UN. Demographic yearbook. http://unstats.un.org/unsd/demographic/products/dyb/dyb_Household/dyb_household.htm, 2013. United Nations Statistics Division.
- [15] UNECE. Private households by household type. http://w3.unece.org/pxweb/dialog/varval.asp?ma=08_GEFHPrivHouse_r&path=../database/STAT/30-GE/02-Families_households/&lang=1&ti=Private+households+by+household+type, 2014. United Nations Economic Commission for Europe Statistical Division Database.
- [16] World Bank. Population, total. <http://data.worldbank.org/indicator/SP.POP.TOTL>, 2015.
- [17] Fernando Pérez and Brian E. Granger. IPython: A system for interactive scientific computing. *Comput. Sci. Eng.*, 9:21–29, 2007. doi:10.1109/MCSE.2007.53.

- [18] Wes McKinney. Data structures for statistical computing in python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 51 – 56, 2010.
- [19] J. D. Hunter. Matplotlib: A 2D graphics environment. *IEEE Comput. Sci. Eng.*, 9(3):90–95, 2007.