

## Twelve-year Follow-up of Respondents in a Sample Survey in Peninsular Malaysia \* (Demographers' Notebook)

By John Haaga

*\* The authors of this note are John Haaga, Director, Committee on Population, National Research Council, United States National Academy of Sciences, Washington, D.C.; Julie Da Vanzo, Senior Economist, and Christine Peterson, senior Quantitative Analyst, RAND, Santa Monica, California; and Tey Nai Peng, Associate Professor, the University of Malaya, Kuala Lumpur, Malaysia. The research reported here was supported by grant No. HD22329 from the National Institute on Child Health and Human Development to RAND. The authors would like to acknowledge with gratitude the comments made on an earlier draft of this note by Shiro Horiuchi, Lynn Karoly and Jacob Klerman as well as acknowledge the excellent work of the National Population and Family Development Board of Malaysia field staff.*

Comparable household-level survey data for the same population at two times are essential for many types of analysis in which age, period and cohort effects must be distinguished. Until recently, it was rare to have usable micro-level data from developing countries collected at two times, far enough apart to allow for significant demographic change in the interim (Hermalin, Entwisle and Myers, 1985). The availability of data sets from the Demographic and Health Surveys (DHS) for many of the same countries in which the World Fertility Survey was carried out has started to improve the situation greatly, enabling cross-national comparisons of changes over a decade or more for countries at different stages of the demographic transition. A growing number of countries have conducted a second DHS survey, and third surveys are planned in some countries, e.g. Indonesia.

Even comparable data from separate samples of the same population at two times, though, are insufficient for some analyses and inefficient for others compared with having data on the same sample at two or more times. The ability to link data on the same families enables, for example, the study of whether fertility expectations are realized over many years, whether and how these expectations change over time, and how decisions about marriage, child-bearing and children's education at one point in time affect family welfare in the future. With panel data, much of what has to be assumed about the linkage of cross-sectional results to the experience of individual families over time can be studied directly, and without the distortions inherent in recall data on many topics. Panel data also enable use of efficient techniques for handling bias owing to unobserved individual-specific or family-specific variables correlated with the variables of interest in a particular model.

Collecting panel data poses significant logistical problems for survey managers. If re-interviews are frequent, the expense of maintaining the sample can be very great. On the other hand, if the second or subsequent contacts are not attempted until years after the first (long enough for meaningful changes to have taken place in the interim), many respondents will be lost to follow-up. It is necessary to look at not only overall follow-up rates but also the selectivity of follow-up. Highly selective attrition from a panel may make the resulting data unusable for some of the research purposes for which they are intended. These problems may not be insoluble, however, and the effort to overcome them may be rewarded by the research value of the data that can be collected.

This note reports the experience of an attempt to find and re-interview in late 1988 and early 1989, as part of the Second Malaysian Family Life Survey (MFLS-2), the female respondents to the 1976-77 Malaysian Family Life Survey (MFLS-1) and a sample of their adult children aged 18 or older. The MFLS-1 female respondents were 1,262 ever-married women (EMW) living in Peninsular Malaysia, aged less than 50 in 1976. The MFLS-1 collected complete retrospective life histories for all the EMW and their husbands, covering marriage, fertility, education, employment and migration. Detailed data on current income, assets and intergenerational transfers were also collected.

The MFLS-1 data set has been widely used by demographers and other social scientists. (A list of publications based on these data is available on request from RAND.) The addition of the MFLS-2 data enables the study of demographic and economic change over the 1976-1988 period for this sample of women and their households.

In this note, we discuss the field methods used to track the panel members and their adult children, report follow-up rates and analyze the selectivity of attrition from the panel, using data from the MFLS-1 on characteristics of both the missing and the re-interviewed respondents and their families. We then discuss the degree to which these results might be generalized to other such attempts at re-contacting survey respondents.

### Field work

MFLS-2 was designed by staff of RAND and the National Population and Family Development Board of Malaysia (NPFDB). Field work was carried out by the NPFDB between August 1988 and January 1989. MFLS-2 conducted interviews with four samples:

- The "Panel" Sample, where the women who had been interviewed for MFLS-1 in 1976, were sought for re-interview 12 years later;
- The "Children" Sample, a sample of the children aged 18 or older of the MFLS-1 respondents;
- The "New" Sample, a new random sample of women aged 18-49 (regardless of marital status) and of younger ever-married women; and
- The "Senior" Sample, a sample of men and women aged 50 and older living in the same enumeration areas as those in the New Sample.

When MFLS-1 was fielded in 1976-77, there were no plans to try to interview the respondents at a later date, and, hence, no attempts had been made to collect any information (national identity card numbers, or names and addresses of close relatives) for later use in tracking the respondents. The NPFDB field scouts and interviewers who traced the MFLS-1 respondents had to rely on names and addresses of respondents as recorded on the cover sheets of the 12-year-old questionnaires. To locate the MFLS-1 primary sampling units (PSUs), MFLS-2 field staff used hand-drawn maps that had been prepared in 1976 by the field staff of Survey Research Malaysia, Sdn. Bhd., a private firm that had carried out the MFLS-1 field work.

In MFLS-2, three teams of scouts, interviewers, data entry staff and their supervisors covered different regions of the country. Field work was carried out in two waves, starting in more remote areas and ending in the Klang Valley (site of the capital, Kuala Lumpur, the main port and several industrial centres), which was the most common destination for interstate migrants. During field work, the teams regularly sent to headquarters lists of names and current locations of respondents whom they learned had moved to another team's territory, so the other team could try to do the interview without waiting for a "mop-up" round.

In both rural and urban areas, the best sources of information for finding MFLS-1 respondents who had moved were the ex-neighbours still living in the original area. One problem that arose fairly often was that these neighbours knew the respondent by a name different from the formal name recorded on the survey forms. In rural areas especially, the village headmen (ketua kampung), who are part-time officials paid by the State Governments, proved to be a useful source of information; many knew where long-time residents had moved. In urban areas, postmen could often remember where respondents had gone. In several cases, the respondent had moved, but kept ownership of the house, and the current tenants were able to supply their landlord's new address. The most difficult PSUs in urban areas were those that contained large blocks of railway worker or government staff quarters: residents move frequently, and people do not know their neighbours as well as in more established settlements.

Urban squatter settlements posed a particular problem. Initially one such settlement could not be located at all from the sketch maps; eventually it was found with the aid of an electricity meter-reader. Further, two such settlements had been demolished for slum clearance.

Many of the respondents of Indian ethnic background were workers on rubber or oil palm estates. Besides ex-neighbours, the best sources included estate clerks (often a worker had been re-assigned to different quarters on a large estate, or to another estate managed by the same company) and shop-owners, some of whom had kept in touch with old customers who still owed them money.

### Follow-up rates for the Panel Sample

There were 1,262 ever-married women who completed interviews in Round 1 of MFLS-1 in 1976. The MFLS-2 interviewers and field scouts learned that 31 of these women had died during the years between MFLS-1 and MFLS-2. Two were reported to have moved out of Peninsular Malaysia, and were dropped from the lists owing to travel and budget constraints. Of the remaining 1,229 women presumed eligible for the MFLS-2 Panel Sample, 889 (72 per cent) were located and successfully re-interviewed with the MFLS-2 Female Life History Questionnaire (MF22). Another 13 of the women were located, but refused to participate in MFLS-2; 21 others were unable to participate for other reasons (most of them because they were never at home during the field work period).

The other 306 presumably moved away; attempts to locate and interview them were unsuccessful. (Some of these 306 may have died or moved outside Peninsular Malaysia. In some of these cases, the MFLS-1 address information was very poor and it was difficult to locate the correct dwelling unit.)

**Table 1: Follow-up of MFLS-1 respondents by ethnic group**

	Malay	Chinese	Indian	Others	Total
Potential (interviewed in MFLS-1, round I)	603	496	148	15	1,262
Reported to have died, 1976-1988	22	7	2	0	31

Reported to have moved outside Peninsular Malaysia	1	0	1	0	2	
	-----	-----	-----	-----	-----	
			-	-	-	
Subtotal: Presumed eligible for MFLS-2 interview	580	489	145	15	1,229	
Completed MF22 questionnaire	(percentage of those eligible) (83.4)	482 (59.9)	293 (69.6)	101 (86.7)	13 (72.3)	889
Unable to locate	87 (15.0)	176 (36.0)	42 (29.0)	1 (6.7)	306 (24.9)	
Refusal	1 (0.2)	11 (2.2)	1 (0.7)	0 (0.0)	13 (1.1)	
Other incomplete	10 (1.6)	9 (1.8)	1 (0.7)	1 (6.7)	21 (1.6)	

Table 1 shows the response rates separately for each of the main ethnic groups in Peninsular Malaysia. Chinese women in the MFLS-1 sample were the least likely to be successfully re-interviewed in MFLS-2 (60 per cent of those presumed eligible), while Malays were the most likely to be re-interviewed (82 per cent). The rate for respondents of Indian background (67 per cent) was between those for the two larger ethnic groups.

### Ethnic and urban-rural differences

As shown in table 2, the differences between ethnic groups were only partly attributable to differences in urbanization (a much higher percentage of Malays than of Chinese live in rural areas). Within each of the strata, follow-up rates were always higher for Malays than Chinese. Within ethnic groups, for Malays and Chinese, follow-up rates were lowest for those who lived in the largest cities in 1976 and highest for rural dwellers. The reverse was true for those of Indian background (the smallest of the three main ethnic groups, accounting for 11 per cent of the population). This reflects the difficulty of tracking estate workers, who comprise the majority of rural Indians.

**Table 2: Follow-up rates for Panel Sample members, by ethnic group and urban/rural residence**

Ethnic group	Residence in 1976 (percent)		
	Metropolitan	Smaller cities	Rural
Malays	63.8	77.2	87.2
Chinese	45.3	65.5	66.8
Indians	80.6	67.7	65.4
Total	54.9	69.8	79.4
Number	(244)	(272)	(713)

### Migration of the respondents

Of those MFLS-1 respondents who were located for MFLS-2 and re-interviewed with the MF22 questionnaire, 63 per cent were found at the same house where they had been in 1976; 28 per cent had moved, but were found in the same district; 4 per cent had moved farther away, but had not crossed a State boundary; and 5 per cent had moved across a State boundary (table 3).

**Table 3: Percentage mobile of MFLS-1 respondents re-interviewed in MFLS-2, by ethnic group, 1976-1988**

Where found in 1988	(compared with 1976 address)	Malay	Chinese	Indian	Others	Total
Same address	63.4	71.3	35.6	61.5	62.8	(559)

Different address within same district	28.4	22.9	44.6	30.8	28.4		
					(253)		
Different district within same State	4.8	2.1	6.9	7.7	4.2		
					(37)		
Different State within Peninsular Malaysia	3.5	3.8	12.9	0.0	4.6		
					(41)		
Total interviewed in MFLS-2	100	100	100	100	100		
	(483)	(293)	(101)	(13)	(890)		
Of those presumed eligible and not living at same address in 1988 as in 1976,		percentage interviewed with MF22 in 1988	68.3	30.9	59.1	71.4	51.1
	(259)	(272)	(110)	(7)	(648)		

The 559 interviewed at the same address as in 1976 constitute 46 per cent of the 1,229 respondents presumed eligible. Of those known or believed to have moved, 331 (51 per cent) were successfully interviewed. Thus, a majority of both movers and stayers were re-interviewed.

To assess whether longer-distance moves are proportionally represented among the Panel women who moved, we compared the 1976-1988 migration of the MFLS-1 respondents who were re-interviewed in MFLS-2 to the 1976-1988 migration rates implied by the migration histories collected for the MFLS-2 New Sample, a representative sample of women aged 18-49 years in 1988. Mobility data for these two samples are presented in table 4. Of Panel respondents aged less than 50 years in 1988 who moved between 1976 and 1988-89 (and it was known where they moved), 9.7 per cent moved to a different district in the same State and 12.7 per cent moved to a different State. However, of New respondents who were ever-married in 1976 and who moved between 1976 and 1988-89, 16.5 per cent moved to a different district in the same State and 30.7 per cent to a different State. Hence, it appears that the follow-up survey disproportionately missed longer-distance movers.

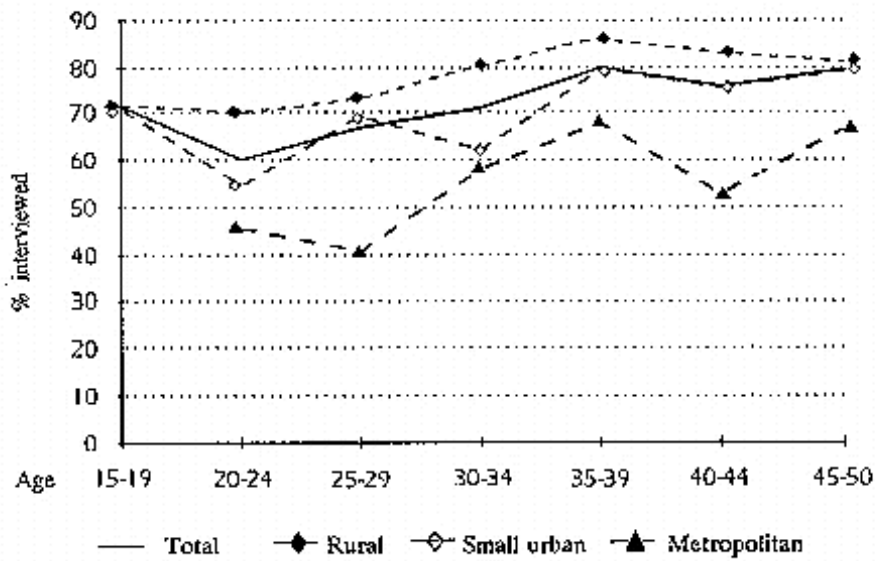
**Table 4: Mobility rates between 1976 and 1988 for comparable sub-samples of MFLS-1 sample and MFLS-2 New Sample, by ethnic group**

Comparison of 1976 and	1988 locations	Ethnic group (per cent)			
	Malay	Chinese	Indian	Other	Total
<b>MFLS-1</b>					
Same address	45.9	42.3	16.9	50	41.1
Different address, same district	27.2	16.1	36.0	30	24.0
Different district, same State	4.8	0.7	3.4	10	3.1
Different State in Peninsular Malaysia	4.8	2.8	7.9	0	3.9
Unable to locate	17.3	38.1	34.8	10	27.4
Total	100	100	100	100	100
	(353)	(286)	(89)	(10)	(738)
<b>MFLS-2 New sample</b>					
Same address	30.6	35.4	23.4	0.0	31.0
Different address, same district	28.8	42.5	53.7	0.0	35.6
Different district, same State	14.5	8.4	9.2	37.5	12.2
Different State in Pensinular Malaysia	26.2	13.7	13.8	62.5	21.3
Total	100	100	100	100	100

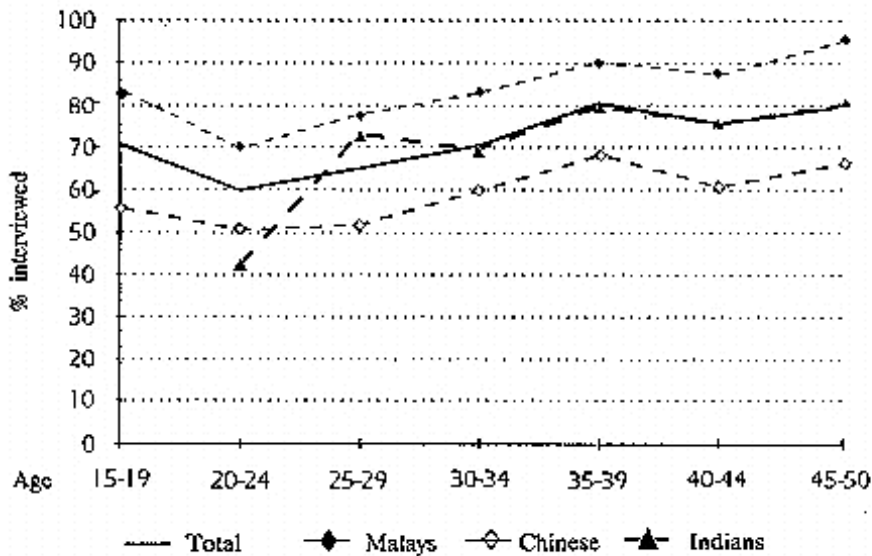
Demographic and socio-economic differentials in follow-up

Overall and in each of the ethnic groups, younger women were the hardest to trace. Figures 1a and 1b show the follow-up rates by age groups, for the total sample, each ethnic group and each rural/urban stratum. Follow-up rates were notably higher for women over age 30.

**Figure 1a: MFLS-2 follow-up rates, by age and urban/rural residence in 1976**

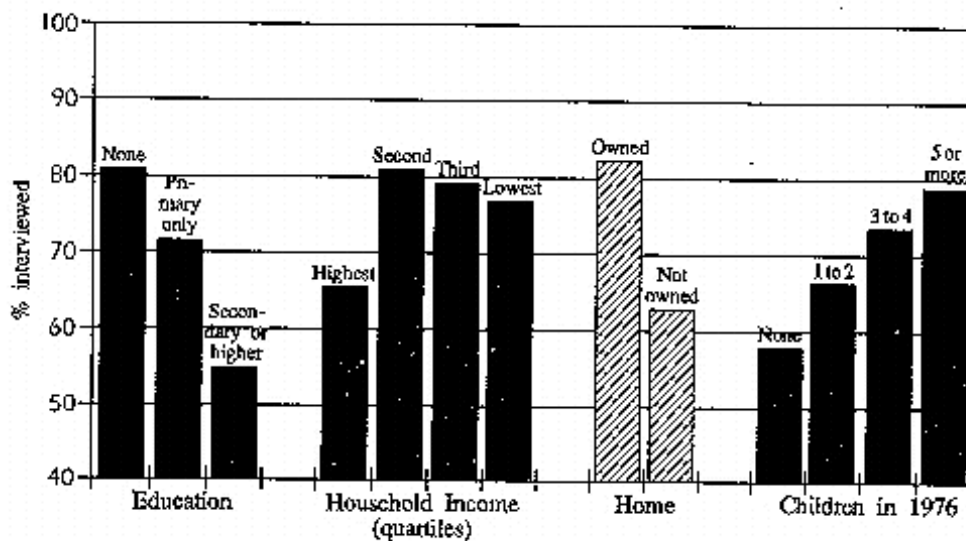


**Figure 1b: MFLS-2 follow-up rates, by age and ethnicity**



Ethnicity, urban/rural residence and age are all associated with educational status, occupation, fertility and household composition in Malaysia. The differences in follow-up rates among ethnic groups and between urban and rural residents are therefore associated with selectivity of the Panel Sample along these other dimensions. In multivariate analyses not reported here, ethnicity, urban/rural residence, age and house ownership in 1976 remain significantly associated with the likelihood of follow-up (Haaga and others, 1993b). Figure 2 shows that the follow-up rates were higher for uneducated women (among whom Malays and rural dwellers were over-represented). The quartile of the sample with the highest household incomes in 1976 had the lowest follow-up rates. Respondents who lived in houses owned by themselves or their husbands were easier to trace than those who were renting their residence or living in relatives' houses in 1976. Women with more children in 1976 generally had higher follow-up rates. One corollary of this is that the loss to follow-up of children or pregnancies covered in the first survey was lower than the loss to follow-up of mothers, though selectivity is obviously still a problem.

**Figure 2: MFLS follow-up rates, by social and economic characteristics**



In MFLS-1, respondents were asked about every time they had changed house since age 15, or their first marriage (whichever was earlier). Table 5 shows follow-up rates for groups of women classified by whether they reported no moves across district boundaries, one, or more than one in MFLS-1 (and classified by age in 1976, as a crude control for differences in length of exposure to the risk of moving). Follow-up rates were indeed lower in each category for those who had moved frequently before 1976. Since people who moved previously are more likely to move again. These results are consistent with those presented previously which show that the MFLS-2 follow-up disproportionately missed (longer-distance) movers.

**Table 5: Follow-up rates for Panel Sample members, by age and prior mobility**

Age in 1976	Percentage moving across district boundaries between age 15 and 49 in 1976, by number of moves		
	None	One	More than one
15-24	69.1	61.0	48.7
25-34	73.2	66.7	61.2
35-49	82.6	81.2	71.3

## Discussion

### Implications of the MFLS-2 for users

The MFLS-2 Panel Sample is a selective subset of the original MFLS-1 sample. Younger, better educated women, those who lived in the largest cities in 1976, those who were more mobile prior to 1976, and those of Chinese background are all under-represented to varying degrees in the Panel Sample, reflecting in large part the higher rates of internal migration and emigration of these groups. But in nearly all the cells of two- and three-way cross-classifications examined here, follow-up rates exceeded 50 per cent – they usually well exceeded that level. Furthermore, for the total sample and for Malays and Indians, the majority of those presumed to have moved were successfully re-interviewed. However, only 31 per cent of the Chinese presumed to have moved were successfully re-interviewed. A number of Chinese may have left the country and thus were not eligible for follow-up. For most of the analyses envisioned, this degree of selective loss to follow-up should not be crippling, provided analysts take it into account both in designing and interpreting studies: for example, by re-weighting to adjust for differential non-response at the second wave (MFLS-2).

Data from MFLS-2 are available for public use from the Inter-University Consortium for Political and Social Research at the University of Michigan, from the National Population and Family Development Board, and from RAND. (For more information about the survey design, see Haaga and others, 1993a. A fuller analysis of the selectivity of follow-up of Panel respondents and of their children can be found in Haaga and others, 1993b.)

### Implications for other follow-up efforts

The experience of the MFLS-2 Panel Sample suggests that fairly high follow-up rates can be attained even after a period as long as 12 years and even when there were no plans for such a follow-up when respondents were first interviewed. In the MFLS-2 Panel Sample, we were able to locate and successfully re-interview over 70 per cent of the women interviewed 12 years earlier in MFLS-1 who were not known to have died or immigrated during the intervening years. The MFLS-2 experience

suggests some recommendations that may help to improve follow-up rates in other panel surveys.

Names of Malaysian Chinese respondents were recorded in MFLS-I only in Roman-letter transcriptions; field workers reported that it would have been easier for them if the Chinese characters for names had been available. Nicknames would have helped for all three ethnic groups. Another suggestion field workers made for future such attempts was that more information about the respondents (such as their own, or their husbands', occupations) be printed on the forms.

Peninsular Malaysia is not very large, so moves within the country do not take migrants too far from their original homes. Migration in Malaysia often takes the form of stage migration, whereby the first urban residence is in a medium-sized city with strong ties to the surrounding countryside rather than in a vast metropolis. Rural-to-urban migrants maintain contact with their places of origin, which helped the follow-up. None the less, even in larger countries, the hinterland of one migration pole is often not larger than Peninsular Malaysia, so that this experience with MFLS-2 could provide a reasonable basis for expectations in other developing countries concerning attempts to re-locate survey samples. As in MFLS-2, former neighbours are likely to be the most important source of information about the current whereabouts of respondents who moved, but many others (e.g. village headmen, shopkeepers) can also provide useful information.

In MFLS-2, residents of large cities were generally the most difficult to find and re-interview. If this experience is typical, then similar follow-up efforts could expect greater success in less urbanized countries. The World Bank has estimated that, around the time of the MFLS-2 survey, 40 per cent of the Malaysian population lived in urban areas: this proportion is relatively low for the group of "lower middle-income countries" within which Malaysia is classified, but roughly the same as for the Philippines and higher than the urban proportions for the neighbouring countries of Thailand and Indonesia (World Bank, 1993).

Follow-up efforts for demographic surveys in the future could be more successful if care were taken to collect information to assist later tracking at the first contact. (For example, in MFLS-2 we collected identity card numbers in case we should later wish to re-interview the respondents.) Addresses are not useful in most developing countries and sketch maps or detailed descriptions are needed to find locations of dwellings. Groups for which higher attrition is expected could be over-sampled in the initial design. During the second wave of field work, attrition rates could be decreased by assigning extra workers, or taking more time, in areas where greater attrition is expected. The scientific value of tracking a panel from an earlier survey, rather than interviewing an independent sample in every survey, would often justify these efforts.

## References

Haaga, J.G., J. DaVanzo, C. Peterson, N.P. Tey, and B.A. Tan (1993a). The Second Malaysian Family Life Survey: Overview and Technical Report, (RAND: Santa Monica, CA), MR-106-NICHD/NIA.

Haaga, J.E., J. DaVanzo, C. Peterson and N.P. Tey (1993b). "Twelve-year Follow-Up of Respondents and Their Children in a Panel Survey in Peninsular Malaysia", Labor and Population Program working paper DRU-380-NICHD, (RAND: Santa Monica CA).

Hermalin, A. I., B. Entwisle and L. G. Myers (1985). "Some lessons from the attempt to retrieve early KAP and fertility surveys", *Population Index*, 51:194-207.

World Bank, World Development Report, 1993, (Oxford University Press: New York).