LSE MIGRATION STUDIES UNIT

Migration Studies Unit Working Papers

No. 2011/08

Scaling Knowledge, Attitudes, Practices and Media Exposure – Are Migrants at Increased Risk of HIV in China?

Paul Bouanchaud

ABSTRACT

HIV, a recent epidemiological phenomenon, is characterised by its complex patterns of spread. Proximately transmitted through specific behaviours of individuals, in theory made more likely because of the particular levels of HIV knowledge and attitudes that those people have, HIV-transmission is also linked with more distal factors, such as socio-economic and demographic determinants, thought have an effect on those levels of knowledge and attitudes and exposure to mass media HIV interventions, and hence people's sexual behaviours. China is a country undergoing significant economic and social change. One effect of this is the rise in levels of rural-urban migration in recent years. Migrants are considered a risk group for HIV, due to their relative levels of marginalisation in their destination cities. This study uses latent trait modelling to develop scales for knowledge, attitudes, sexual behaviours and media exposure from a survey performed in six Chinese cities in order to investigate how well the survey items measure what they are supposed to. The scales are then used in further analyses to investigate whether migrants and non-migrants have different levels of HIV-related knowledge, attitudes and behaviours, and whether these factors have an influence on condom use. The results suggest that contrary to expectations, migrants out-perform non-migrants on a number of the scales, although they do appear to have significantly more stigmatising attitudes and poorer levels of media exposure. The analysis also suggests that levels of HIV-avoiding knowledge and reported behaviours, and exposure to the media are all significant predictors of condom use, but that levels of stigmatising attitudes are not. Furthermore, once these variables have been controlled-for, migrants' levels of condom use are no different to their non-migrating counterparts. The study makes a number of recommendations, including the alteration of particular survey items and the development of more targeted HIV interventions to reduce migrant's levels of stigmatising attitudes, and calls for more research to be done both into the development of scaling in HIV attitudinal and behavioural surveys and into the factors that drive Chinese migrants' attitudes towards HIV.

Acknowledgements

Abstract

HIV, a recent epidemiological phenomenon, is characterised by its complex patterns of spread. Proximately transmitted through specific behaviours of individuals, in theory made more likely because of the particular levels of HIV knowledge and attitudes that those people have, HIV-transmission is also linked with more distal factors, such as socio-economic and demographic determinants, thought have an effect on those levels of knowledge and attitudes and exposure to mass media HIV interventions, and hence people's sexual behaviours.

China is a country undergoing significant economic and social change. One effect of this is the rise in levels of rural-urban migration in recent years. Migrants are considered a risk group for HIV, due to their relative levels of marginalisation in their destination cities.

This study uses latent trait modelling to develop scales for knowledge, attitudes, sexual behaviours and media exposure from a survey performed in six Chinese cities in order to investigate how well the survey items measure what they are supposed to. The scales are then used in further analyses to investigate whether migrants and non-migrants have different levels of HIV-related knowledge, attitudes and behaviours, and whether these factors have an influence on condom use.

The results suggest that contrary to expectations, migrants out-perform non-migrants on a number of the scales, although they do appear to have significantly more stigmatising attitudes and poorer levels of media exposure. The analysis also suggests that levels of HIV-avoiding knowledge and reported behaviours, and exposure to the media are all significant predictors of condom use, but that levels of stigmatising attitudes are not. Furthermore, once these variables have been controlled-for, migrants' levels of condom use are no different to their non-migrating counterparts.

The study makes a number of recommendations, including the alteration of particular survey items and the development of more targeted HIV interventions to reduce migrant's levels of stigmatising attitudes, and calls for more research to be done both into the development of scaling in HIV attitudinal and behavioural surveys and into the factors that drive Chinese migrants' attitudes towards HIV.

Table of Contents

1 Introduction

HIV is a relatively recent phenomenon in global epidemiology, and has been termed a "disease of post-modernity" (Kallings, 2008), recognising the complexity and local variability in the disease's interpretation by individuals, and their concomitant behaviours. Proximately caused through having unprotected sex with an infected person, sharing needles or blood products or during childbirth or breastfeeding, but more distally related to other social, demographic, economic and psychological determinants, HIV is a heterogeneous social phenomenon. Whether someone knows about HIV, and possible ways to avoid contracting the virus seems perhaps to be an obvious factor, but this apparent simplicity belies a more complex reality, where knowledge in turn is determined by a host of factors, and what each individual does with their knowledge is not necessarily always predictable.

China is experiencing unprecedented levels of rural-to-urban migration, estimated to have reached more than 150 million people in 2007 (Chan, 2008). While there has been much research into the correlation of population mobility and HIV prevalence, particularly in sub-Saharan Africa, there remains a gap in the literature regarding the analysis of intermediate variables, such as socio-economic factors and their impacts on particular attitudes and behaviours (White 2003), and this dissertation aims to go some way towards addressing this gap in the Chinese context.

The orthodox view is that migrants have a higher risk of contracting HIV and of transmitting the virus when they return home. Several theories suggest that migrants are more likely to engage in risky behaviours (Yang, 2007; Brockerhof & Biddlecom, 1999) and so be a greater risk of HIV infection. Previous quantitative studies in China connecting migrants' levels of knowledge, attitudes behaviours and practices (KAB/P) in relation to HIV have tended to rely descriptive statistics, failing to account for potentially confounding variables (Sheeran & Taylor, 1999). The aims of this study are twofold: first, to apply a latent trait analysis to a survey that investigated levels of KAB/P and media exposure among different groups in China, in order to gain insights into how well the standard battery of questions function as effective measures in this context, and second, to interrogate the hypothesis that migrants are at greater risk of HIV in China.

1.2 Outline of the Dissertation

Section 2 provides a discussion of the context of the HIV situation in China – its history and the groups that are considered at most risk of contracting the virus. There follows a discussion of the drivers of rural-urban migration in China, and how this 'floating population' might be considered to be one of the key risk groups for HIV. The role that the media might play in disseminating HIV messages is then briefly explored, before a

discussion of previous approaches that have been taken to analysing HIV survey data, and the justifications for the method applied here are presented. The research questions for the study, which draw from the themes outlined in Section 2 are then listed. Section 3 describes the survey dataset and methods used in its analysis. Section 4 presents the results of each step in the analysis accompanied by a discussion of their implications with reference to the particular research questions of this study, alongside some criticisms of the study and discussion of potential weaknesses in the findings alongside some potential avenues for future work. Finally, Section 5 tentatively draws some conclusions from the various findings of this paper.

2 Context

2.1 History of China's HIV Epidemic

China's HIV/AIDS epidemic has gone through three distinct phases to date (Jie et al, 2004). The "entry phase", from 1985 to 1989 was a period with very few reported cases of HIV, with most being in foreign nationals and in a small number of haemophiliac patients (contracting the virus from imported blood products) in coastal urban areas (ibid.). The "spreading phase", 1989 to 1994, saw HIV infections being passed among distinct communities, still in relatively low numbers, (for example among 146 people in a city in Yunnan province), among small groups of intravenous drug users (IDU's), and among limited numbers of commercial sex workers (CSW's) and sexually transmitted infection (STI) patients (ibid.). The third phase, characterised by an expansion of the epidemic began in 1994 and continues to the present. This period has seen large numbers of IDU's become infected in Yunnan, with spread to other provinces, along with large numbers of blood donors in central China being found to be carrying the virus (Guomei, 2005). There is also evidence of increasing numbers of STI patients and CSW's testing HIV positive, and reported rises in heterosexual transmission rates in recent years (ibid.).

By the start of 2008, there were an estimated 700 000 people living with HIV/AIDS (PLWHA) in China (UNAIDS 2008), the vast majority of whom are not aware that they are carrying the virus (Jie et al. 2004). Just less than half are thought to have contracted HIV through IDU, with a similar number having acquired it through unprotected sex (UNAIDS, 2008). The epidemic in China, however, remains characterised largely by its continued focus in specific high-risk groups. As in many other national epidemics, these include IDU's, CSW's, men5 who have sex with men, and migrant populations (ibid.). Of particular concern is the possibility that some of these groups may become "bridge" populations, allowing large-scale spread of HIV infection into the wider Chinese population (Qian et al., 2004). There is some evidence that this has begun to occur. Data from Yunnan Province also show a shift towards heterosexual transmission, with just 5.3% of infections in 1995 being through this route, rising to 11.4% ten years later (Xiao et al., 2006), and from 5.5% to 10.9% in Jiangsu Province for the same period (Haitao et al., 2004). Shifts in the ratio of men to women infected with HIV also suggest that the virus is spreading outside its original risk groups. In Yunnan in 1990, there were 40 men to every woman infected. By 2004, this had fallen to 3.8 to 1 (ibid.), indicating that men have started to infect their partners, many of whom are not able to negotiate safe sex due to gender power differentials, particularly in poorer or more rural areas (Guomei, 2005). It also seems likely that the virus has spread into the female population through their increasing rates of drug use, and rising numbers of female CSW's becoming infected with HIV (ibid.).

2.2 Government Policy

A country's capacity to cope with the HIV epidemic depends on the levels of political commitment and the state of its health care systems (Grusky et al., 2002). In 1998, ix years after organising its first national AIDS conference, the Chinese government produced its first Long- and Medium Term Plan for HIV/AIDS. In the same year, the government set up a fund for the prevention and control of HIV, which in recent years has seen large increases in investment (ibid.). The government also set up sentinel surveillance sites around the country (Shao & Lv, 2004), implemented blood screening and condom promotion (Grusky et al., 2002). However, there was little public acknowledgement of the seriousness of the situation by high-ranking government officials until a 2001 speech by the deputy health minister (ibid.). Recognising the importance of HIV as a potential threat to security as well as health, in 2003, the government launched the "Four Frees and One Care" policy[1] which marked a significant scaling-up of efforts to combat the disease (Shao, 2006). However, despite the political will, bottlenecks in the system remain, acting as barriers to full policy implementation (ibid.). Calls have been made for more evidence-based treatment and prevention models that are suitable to China's changing social and economic situation to be developed, in particular in behavioural research, which appears somewhat lacking (Haitao et al., 2004). Another barrier to effective programmes in China could be the relative weakness of its "civil society" sector (Kaufman & Jing, 2002). Worldwide, many successful HIV prevention programmes have been driven by civil society and NGO's (Epstein, 2007), and while there have been some advances in levels of NGO and quasi-NGO involvement in China in recent years, policy and program development remain almost exclusively in the hands of central government (Kaufman & Jing, 2002). This may hinder localised, small-scale community-led responses to the epidemic, which may be more able to focus on particular groups at risk.

2.3 Risk Groups

There are features of particular groups of people in China, as elsewhere, that make them more vulnerable than others to contracting and transmitting HIV. There are estimated to be 3.5 million Intravenous drug users (IDU's) in China (Borowitz et al., 2004), and this group appear to have been the principle locus of HIV spread in the earlier stages of the HIV epidemic in China. Yunnan province, part of the Golden Triangle of drug production and distribution, (Garten et al., 2004) has some of the highest IDU rates in China, and there is evidence of a shift in the last 20 or so years from inhalational to injecting drug use (ibid.), alongside high levels of needle-sharing among IDU's – over 70% in one study in Yunnan (Xiao et al., 2006). The province was one of the first to experience high rates of HIV among its IDU's, probably brought in to the area through cross-border drug trafficking from neighbouring countries (Guomei, 2005). This hypothesis is supported by

---

1   Free treatment, voluntary counselling, prevention of mother to child transmission and schooling for AIDS orphans.

molecular epidemiological evidence that suggests that the HIV subtypes in Yunnan, mostly clade B recombinants, were most closely related to those circulating in Laos, Myanmar and Thailand during the period (Garten, et al., 2004; Graf et al., 1998). The epidemic has spread beyond the borders of Yunnan among IDU's and is infection rates among these groups have been found to be significant in a number of other areas. There is also evidence that IDU's are more likely to have multiple sexual partners, which may amplify the potential for HIV-spread (Xiao et al., 2006). Commercial sex workers (CSW) are also a group considered at high risk for HIV. The sex trade was largely eradicated following the establishment of the People's Republic in the 1950's, but since the beginning of the 1980's, rising levels of wealth in the cities, and increases in the number of female rural-urban migrants have both led to CSW becoming much more common (Jie et al., 2004), with current numbers estimated to be above 3 million (Borowitz et al., 2004). Characteristically, CSW's have low reported levels of consistent condom use (Hong et al., 2009) and relatively high levels of sexually transmitted infections (ibid.), both of which are considered to increase the risk of contracting HIV. Men who have sex with men are also thought to be at increased risk of HIV in China, as elsewhere, due to higher numbers of sexual partners, higher frequencies of multiple-partner sex and the underground nature of their sexual networks in China, where traditionally such practices are frowned-upon (Yinhe, 2006; Yu-Hua et al., 2004).

Finally, migrants are a group that has been found to be at higher risk of HIV infection (Qian et al., 2004; Yang & Xia, 2006; Yang et al., 2007). This is thought to be both directly because they are migrants (see section 2.4 below), and because as migrants, they may be more likely to become members of the other risk groups mentioned above. Where two or more risk groups have substantial crossover, there can be a greater chance of HIV being communicated through these "bridges" to infect other people or networks of people (Haitao et al., 2004). Several studies have found that migrants have lower levels of knowledge (Rongyao et al., 2000; WHO, 2005) and poorer health-protective sexual behaviours in China (ibid.). Female migrants have been found to be particularly vulnerable to engaging in commercial sex, with rates between 14 and 80 times those of non-migrant women (Yang & Xia, 2008; Yang & Xia, 2006); Yang et al., (2005) reported that 6.3% of female migrants in their study had ever exchanged sex for money. There is also evidence that male migrants are more likely than non-migrants to engage in IDU (ibid.), and male unemployed migrants have been found to be more likely to be blood donors (Qian et al., 2004); migrant blood donors being found to have significantly higher rates of HIV infection than non-blood donors in the same study. Furthermore, migrants have been found to have higher rates of multiple sexual partners – 28% in one Yunnan study (Xiao et al., 2006), they have been shown to have poor, but increasing levels of knowledge about HIV (Zhou et al., 2007), and relatively low rates of STI and HIV testing and willingness to be tested (Wang et al., 2008). Considering migrant's apparent increased risk, and crossovers with other risk groups, the unprecedented migration in

China, exceeding 100 million for the first time in 2004, and estimated to be approximately 150 million people today, (Chan, 2008) presents a serious challenge to controlling the spread of HIV in their origin and destination points.

2.4 Migration

2.4.1 A Working Definition

Migration is considered a difficult phenomenon to define and quantify in the field of demography, and this makes its analysis in many contexts conditional on whichever definition is chosen. Within the Chinese context however, migration appears to be a more clearly delineated concept. The Chinese government's use of residence permits (*hukou*) allows a straightforward definition of migrants - being people who live in a city without having a *hukou* document for that city, as is common practice in previous work on Chinese internal migrants (Zhou et al., 2007; Rowland, 1992). It is to this "floating population" of workers that I am referring henceforth in discussions of migrants in China.

2.4.2 Economic Development: A Driving Force

Following the market liberalisations that began to occur in the late 1970's, China has undergone huge social and economic changes that have had wide-reaching effects throughout the country, including large-scale rural-urban migration. A major benefit has been the substantial rises in average incomes, with official estimates indicating that between 1978 and 2005, over 200 million fewer rural people were living in poverty – a fall in the rate from 30.7% to 2.6% (Wang & Cai, 2007) – if true, an achievement of impressive proportions. It might be questioned, however whether part of this reduction is due to the huge migrant flows out of the countryside – simply moving some of the poor elsewhere, and increasing the per-person amount of land available for farming (Wang & Cai, 2007).

2.4.3 Income Inequality

Possibly supporting this conjecture, the number of urban poor in China, fairly negligible until the 1990's, has seen large increases in recent years; one estimate indicating a 20-fold increase in numbers living in poverty in urban areas between 1990 and 1999 (Murphy 2006). This is in part due to macroeconomic changes in labour markets with shifts from a centrally-controlled economy leading to rises in unemployment, but has also been attributed to the large inflows of migrant workers from the rural hinterlands (ibid.). Increased urban poverty has been accompanied by rising income inequality, both intra- and inter-provincially: both a driver and consequence of urban-rural migration. Migrants moving to cities tend to have lower levels of education and thus employment opportunities tend to be limited lower-paid work (Dutton, 1998). As Murphy (2006) notes, rural-urban migrants tend to occupy relatively precarious and poor living conditions in urban areas, and

tend to be employed in more dangerous and dirty work than their urban-originating counterparts. This has created large disparities between groups in urban areas (Laquian, 1996). There is also evidence of city-dwellers having growing levels of suspicion towards incoming migrants. The term *liumang* (流氓), equivalent to 'hooligan', is made up of the characters *liu* (to flow) and *mang* (used to describe people from 'beyond the boundaries'), is described by Bakken (2000) as the way many urbanites think of the outsiders; outsiders here as Dutton (1998) contends, not just spatially, but socially as well. Poverty levels have been found to be approximately double among rural-urban migrants compared with the registered urban residents (Li, 2001). This places non-*hukou* holding urban dwellers in a potential situation of vulnerability, both due to their material circumstances and in terms of discrimination they may face in the cites.

Two large-scale Chinese studies have shown that through remittances, migrants are able to increase the wealth of their families in rural areas (Li, 2001, Li 1999), but this can create income divides between families of migrants and non-migrants families in rural areas (Murphy, 2006). The negative effects of higher levels of income inequality on health outcomes are well established in the literature (see Wilkinson & Pickett, 2005 for a review), and in China in particular (Zhongwei, 2006). Furthermore, many studies in developing countries have indicated that income inequality specifically can be a major macro-level driver for the HIV epidemic (Gillespie et al., 2007; Barnett & Whiteside, 2006). This raises interesting questions about the role that migration might play at a structural level, in raising the risks of HIV spread in both urban and rural areas of China.

2.4.4 Why Are Migrants at Risk?

As Yang (2006) notes, most studies on HIV in China tend to view migrants simply as vectors for disease transmission without attempting to formulate a theoretical framework for why this should be so. He argues for a move away from this reductionist viewpoint, saying that migration should be seen as a root cause of HIV risk behaviours, with the economic and social milieus of post-migration needing to take a more central role in analyses. "Round trips between villages and cities are not *per se* a vector for HIV: the key issue is the *behaviours* of migrants" (Lagarde et al., 2003). In order to ground this viewpoint, the following theoretical perspectives go some way to explaining why migrants might be at greater risk of HIV: Social control theory states that people behave in ways that are socially sanctioned, and conform to societal norms because they fear the consequences of not doing so. If people are separated from close family and peers, and the social control structure therein, they are more free to behave differently (Stack, 1994; Yang et al., 2007; Barnett & Whiteside, 2006). Social isolation theory (Wilson, 1987) can make another important contribution to our understanding of why migrants might behave differently. When

people are isolated from mainstream society, and feel that they are being denied full access to the opportunities that the mainstream have, then they are less likely to conform to the norms of that society. This in turn makes them more likely to engage in behaviours that might be considered deviant by the mainstream (Yang et al., 2007). Both of these theories may apply to the situation of migrants in China, going some way to explaining why they might engage in more epidemiologically risky behaviours. And besides these theories, migrant behaviour may be less a matter of choice than necessity, their social and economic marginalisation potentially meaning that they are more easily coerced into sex work (Yang et al., 2005).

Migrants have also been ascribed certain peculiar psychological characteristics that may make them more likely to engage in riskier behaviours. The self-selection of migrants means that they might be more adventurous or willing to take risks than their non-migrating counterparts (Brockerhof & Biddlecom, 1999). Furthermore, economic migrants in China tend to be predominantly young single adult males, a group who are thought to behave in more risky ways than others (Yang et al., 2007).

2.5 Knowledge, Attitudes, behaviours and practices (KAB/P) studies

2.5.1 A Theoretical Framework

Intervention strategies have largely focussed on ensuring that people at risk of HIV have the necessary "tools" at their disposal to be able to make sound judgements about risks and safer sexual behaviours. "Sound knowledge about HIV is an essential pre-requisite – albeit, often an insufficient condition – for adoption of behaviours that reduce the risk of HIV transmission." (UNAIDS, 2007:52). Measuring the levels of a population's knowledge about HIV transmission is considered important in understanding to what extent the epidemic might be avoided (or conversely, spread) in that population, and the magnitude of the challenges that intervention strategies may face (Chen et al., 2003).

The link between people's attitudes and knowledge of HIV and their subsequent behaviours is therefore one of the central themes in current strategies to limit the effects of the epidemic (UNAIDS, 2007). The theoretical basis for this study rests on a number of previous attempts to understand the relations between individual beliefs, knowledge, etc., and their subsequent behaviours. First, the information provision model, as outlined by Fisher & Fisher, (1992) (in Sheeran & Taylor, 1999) simply states that information about

HIV and possible ways to avoid it are prerequisites for adopting health-protective practices. Second, the health belief model (Becker et al., 1977), is also useful, whereby an individual's subjective beliefs about the likelihood of certain health outcomes are considered to be determinants of their behaviours. These feelings are thought to be informed largely, although not exclusively by the individual's social and demographic characteristics, taking effect through socialisation. With reference to condom use, the beliefs would be concerned with a person's understanding of the dangers that HIV poses, and their beliefs surrounding the practice of protective behaviours (Sheeran & Taylor, 1999).

Finally, Ajzen and Fishbein's (1977) theory of reasoned action, describing the relations between attitudes and behaviours may be relevant. This model has more recently been applied to HIV prevention strategies specifically (Fishbein, 2000) and serves as another point from which to understand what the determinants of particular behaviours are likely to be, along with factors that might act as mediators, either hindering or encouraging more health-protective behaviours in individuals. Fishbein (2000) argues that besides knowing that a particular behaviour is protective against HIV, people must also have the intention to perform that behaviour, and that intentions to perform particular behaviours are informed by attitudes towards the behaviour, perceptions of norms surrounding its performance, and perceptions of one's self-efficacy in achieving the behaviour. This approach highlights the multidimensional nature of behavioural decision-making, something that KAB/P studies try to capture through the range of topic areas covered. A meta-analysis of various psychological and socio-economic determinants effects on planned condom use (Sheeran & Taylor, 1999) further supports the inclusion of these variables in research into safer sex practices. The study showed that knowledge about HIV, and various demographic indicators were correlated with condom use, and that in particular, previous safer-sex practices and positive attitudes to condom use were very good predictors of condom use. The authors of the meta-analysis note, however that the majority of studies they analysed did not control for confounders, meaning that some of the effects may have been obscured or inflated in the underlying mechanisms. The present study aims to partially address this gap in the literature through the use of regression analyses on various behavioural, attitudinal and demographic variables in relation to condom use, which controls for some of these potential confounders.

2.5.2 Theoretical Framework for Migrant HIV Risk

Utilising the conceptual approaches presented by Smith (2005) and Gorbach, et al (2002), a unified framework is presented here depicting the hypothesised relationships between the many intermediating factors that are considered to have significant influence on HIV risk, taking the

theoretical positions outlined above. The green boxes show the variables of particular interest to this study; the red boxes indicate the theoretical influences that are hypothesised to play a role in increasing migrants' preponderance to risk taking. The dashed arrow indicates the two specific hypothesised connections that will be tested later: that there is something peculiar to migrants alongside their usually lower levels of KAB/P, and that two these things place them at increased risk of HIV; and that different levels of KAB/P influence levels of epidemiologically risky behaviours, here using the measure of "ever used condoms", a crude indicator, but one that is frequently used in the literature (Sheeran and Taylor, 1999).

Figure 2.1: Theoretical framework conceptualising HIV risk and migration.
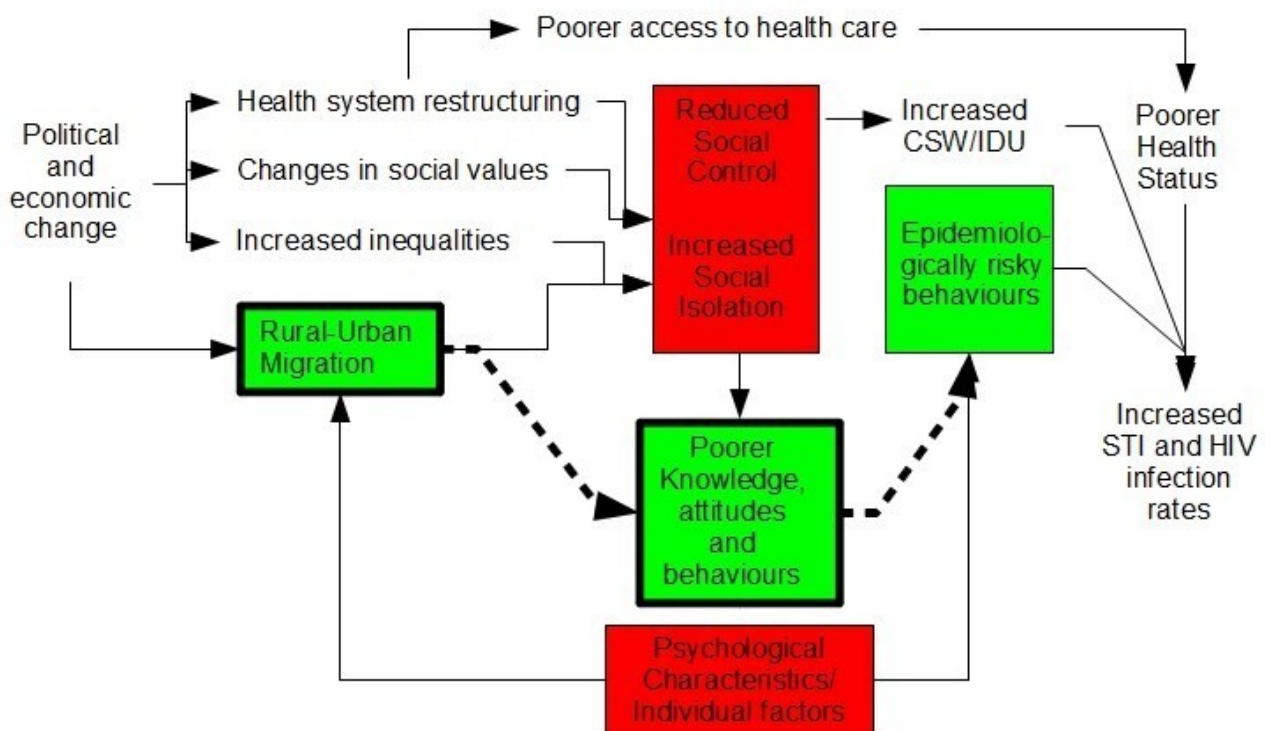
2.5.3 Media and KAB/P

The provision of information is broadly recognised as a major feature of HIV prevention, care and support worldwide, and mass media communication of this information is thought to be an effective means of improving people's knowledge, attitudes, and health-protective behaviours. Being able to assess levels of media exposure is an important consideration in the planning of HIV information campaigns (Shisana & Simbayi, 2002). The use of mass media, however runs the risk of homogenising its target audience, and it is for this reason that it is important to understand how different groups are exposed to the media, to assist in the targeting of information campaigns. Shisana and Simbayi (2002) note however that while it is not possible (nor indeed desirable) to reduce changes in KAB/P to specific interventions such as mass media

campaigns, any reductions that are brought about in risky behaviours or negative attitudes are inevitably the product of communication of one form or another, be it interpersonal, community-level, or mass media-based, and so begin able to quantify even one of these is a good target for research.

2.5.4 Previous Approaches to KAB/P Studies

KAB/P studies are a well established means of obtaining data on the factors that can act as barriers to HIV prevention and care. The surveys are well validated and relate to key areas of the UNGASS guidelines on Monitoring the Declaration of Commitment on HIV/AIDS (UNAIDS, 2005). The majority of previous analyses of HIV KAB/P studies have centred on either purely descriptive statistics (e.g. Zhou et al., 2007) or classical test theory (CTT) approaches, using sum scores.[2] While these summed composite scores are easily calculated, this paper will argue that more statistically powerful techniques might be appropriate. Some disadvantages of the CTT approach include assuming that all of the questions have equal weight in determining a person's knowledge, attitudes, etc. about HIV, and assuming that all of the questions are asking what they are designed to ask. This paper aims to go some way towards assessing whether these assumptions can indeed be made in the case of HIV knowledge, attitudes and practices in China.

Several studies have utilised statistical methods such as factor analysis to generate scales/indices for scoring individual respondents in relation to HIV (Examples include: Gerbert et al., 1998; Darke et al., 1991; Kalichman & Rompa, 1995; Berger et al., 2001; Holtzman et al., 2003; Wang et al., 2007; Lönn et al., 2007). Factor analysis is a familiar technique to many researchers and is part of a family of statistical models using item response theory (IRT). IRT permits a large number of items to be aggregated into an index, with the advantage of different weights being applied according to the distributions of the items on a latent scale. This method also allows for individuals scores to be calculated, and then used in further analyses. However, as will be discussed later, factor analysis has a number of potential drawbacks in relation to its use in survey data analysis. Therefore, an alternative, but functionally equivalent method must be used to develop weighted scales. Here, as there are large numbers of binary response variables in the dataset, and we are assuming that scales of knowledge, attitudes, behaviours and media exposure are all continuous, latent trait modelling is the most appropriate method to develop KAB/P and media indices (Bartholomew & Knott, 1999).

While the literature is extensive on the use of latent trait models for the development of scales in the fields of psychology, education and sociology, their use in public health and epidemiology remains rare. More

---

2   Sum scores in KAB/P-type surveys are the composite scores for each individual in each area of the survey.

specifically in sexual risk behaviour analysis, latent trait models are virtually unknown (Mattson et al., 2008). A search of the HIV literature has shown that, latent trait analysis does not appear to have been used for KAB/P-type surveys to date, although there is evidence that it is beginning to gain currency among other sexual risk researchers, where the modelling of the correlates of different behaviours might be useful (Fendrich et al., 2008), and where weighting of relatively rare question responses is important (Hennessey et al., 2008). More particularly, the application of latent trait analysis to self-reported health behaviours (such as those in a KAB/P-type study) remains under-explored, (notable exceptions include: Mattson et al., 2008; Rao et al., 2006; Hennessey et al., 2008 and McClelland et al., 2002), although its use in other forms of self-reported behaviours is more widespread (for example, Osgood et al., 2002). Advantages of IRT over CTT include: being able to analyse each survey question's performance (see section 3.2 below), being able to get scores for individuals that can be used in other analyses, having scales may be developed simultaneously from a range of different types of data (ordinal, continuous, binary), and having a robust way to cope with missing values (Mattson et al., 2008).

The results of the survey analysed in this report lend themselves well to latent trait modelling for a number of reasons. The survey is designed to explicitly focus on four areas – knowledge, attitudes and beliefs/practices and media exposure in relation to HIV, and as such, substantively these areas may be seen as latent variables (or sets of latent variables), each not directly measurable, but partially revealed through individuals' responses to the battery of questions in the survey. Furthermore, the survey questions are in a number of formats, from binary 'yes/no' options to Likert-type scales, and continuous variables.

2.6 Research Questions

Themes from previous analysis of KAB/P-type data and theories outlined above suggests that rural-urban migrants are likely to be at an increased risk of HIV both due to their particular circumstances directly and because of their lower levels of health-protective behaviours and knowledge. The following research questions attempt to operationalise these relationships, and aim to contribute to the literature, both in terms of the methodological approaches that may be taken in KAB/P analysis, and the substantive results relating to migrants' HIV risks in China:

- How well do KAB/P survey items function in measuring what they are designed to measure,

and to what extent is latent trait an improvement on sum scoring?

- To what extent do migrants' KAB/P and media exposure differ from those of non-migrants?
- What effect do the KAB/P and media exposure levels in the sample have on individual likelihood of using condoms?
- Does the theoretical framework hold with regards to migrants? Are there intermediate variables that play a role in their HIV risk, such as socio-economic variables, or are migrants inherently psychologically different to non-migrants?

3 Methods

This section describes the dataset used in the analyses and then goes on to delineate the operationalisation of the research questions given in section 2.6. Greater attention is paid to the details of the latent trait modelling, as this technique is less commonly found in the analysis of survey data of this type. The linear and logistic regression models are then outlined.

3.1 Data

The China AIDS Media Partnership (CHAMP) 2008 KAB/P survey was carried out in six Chinese cities (Beijing, Kunming, Shanghai, Shenzen, Wuhan and Zhengzhou) between February and March 2008 by Renmin University with the aim of describing the levels of HIV/AIDS-related knowledge, attitudes and behaviours/practices (KAB/P) and exposure to various forms of media among four key groups of people – blue and white collar workers, migrants and youths (CHAMP/UNAIDS 2008). Here, as migrants are the key focus, these groups have been dichotomised into 'migrants' and 'non-migrants'. A complete description of the sampling frame and interview procedures may be found in CHAMP/UNAIDS (2008). The report that was published contains detailed descriptive statistics exploring the various levels of KAB/P and levels of media exposure, but to my knowledge, no attempt has been made to use statistical modelling on the dataset in order to generate a more nuanced picture of the various interrelated factors involved in influencing the respondents' reported KAB/P's and media exposure levels.

3.2 Latent Trait Modelling

Latent variable models in the broadest sense are essentially examples of generalised linear models, but with latent rather than observed independent variables (Bartholomew & Knott, 1999; Stares, 2008), whereby a larger number of manifest response variables are linked through regression models to a smaller number of (unobserved) latent explanatory variables (ibid.). The latent and manifest (observed) variables in the model may be categorical or continuous, and the choice of model used depends on which combination of these types of variables is present. The most commonly encountered type of latent variable model, factor analysis, is intended to be used when both latent and manifest variables are continuous (Stares, 2008; Bartholomew et al., 2008), however further model types have been developed to cope with other combinations of variable types.

Since the majority of items concerning KAB/P in the survey analysed here are categorical (some binary, others in a Likert-type scale format), latent trait analysis is the best-suited method for analysis (Bartholomew et al., 2008). As in regression models in which there are binary response variables, where it would be inappropriate to use a linear link function, here a logarithmic function is used to estimate the odds of

particular responses being made. The reasons for needing to use a logarithmic function are outlined below, but it is worth noting that they are frequently ignored in survey data analysis, where factor analysis is often used, treating categorical items as continuous (Mattson et al., 2008; Stares, 2008).

Using factor analysis for categorical variables is problematic for a number of reasons, and as Stares (2008) notes, is analogous to applying a linear regression model to categorical dependant variables. In a situation in which the response (y) variables are binary, estimating the expected values of y is equivalent to modelling the probability of getting either a 'yes' or 'no' response. Since probabilities axiomatically must lie between 0 and 1, using a linear function to relate the latent (x) and manifest (y) variables means that there is a possibility of estimating probabilities that lie outside of the 0 - 1 range. This would render the results meaningless. Furthermore, it is quite possible that the rate of change in the probability of a particular response might be different for different levels of x, in which case a curvilinear relationship might be more appropriate. In order to avoid these problems, a non-linear function linking x and y is used (Bartholomew et al., 2008). The definitions of the models used in this analysis are given in more detail in Appendix 1.

3.2.1 Interpreting Latent Trait Parameter Estimates

For each observed item i the analysis models two parameters of interest to this study. The intercept, α_{i0}, is also known as the "difficulty parameter", and concerns the probability of the median individual giving a "correct" response (Bartholomew et al., 2008) (for example in the binary items in the survey, answering a question relating to factual knowledge about HIV transmission routes correctly, and getting a '1' score). These are useful in understanding a latent variable's meaning, but also play a role in informing future HIV survey design: for example, if a survey item on HIV transmission knowledge is found to have a very low difficulty score, with a large proportion of respondents answering it correctly, then that item is rather ineffectual in defining the knowledge scale, and so could be replaced in future surveys, or ignored in scaling HIV knowledge.

The second set of parameters of interest are the discrimination parameters (α_{ij}). These are the factor loadings, with a larger value showing a greater effect of factor j on the probability of responding correctly to item i. In other words, the higher the value of α_{ij} for and item, the bigger the effect of the latent trait on the odds of giving a particular answer to the survey item. Thus that item would make it easier to discriminate between two people on the scale. Graphically, the α_{ij}'s estimate the slope of the item characteristic curve (ICC), as we shall see later. As with the difficulty parameters, items showing high discrimination power (a steeper curve) are useful in in differentiating between

individuals on the latent dimension, and so these items are doing more 'work' (Stares, 2008) and should be kept in future surveys and scales.

ICC's are graphical representations of fitted probabilities of item responses, combining both difficulty and discrimination parameters for a range of values for a latent trait x_j. They permit us to see how the probability of choosing a particular response category changes at different points along the latent trait. Along with the slope of the curve – the item discrimination, the horizontal location of the curve in the plot is defined by the difficulty parameter(s), further along the latent trait meaning a more "difficult" question.

The latent trait estimations were performed using the Latent Gold 4.0 software package (Vermunt & Magidson, 2005), using full information maximum likelihood estimation, which has been shown to robustly cope with missing values (Schafer & Graham, 2002), and the software default values (see Vermunt & Magidson, 2005). Before running the analyses, binary variables regarding factual questions in the dataset (knowledge of correct ways to avoid HIV infection, for example) were recoded so that a correct answer was coded 1, and incorrect and "don't know" answers were coded 0. This was to aid interpretation of the item's functioning later, and the recoding of "don't know" items follows the recommendations of UNAIDS (2007). Also items with Likert-type response options were recoded into binary items, with the exception of those relating to attitudes which were recoded to have three possible responses (positive, negative and neutral attitude). ICC's were produced using a function written by Dr Jouni Kuha for the S-Plus software package (S-Plus, 2006). Posterior scores on each trait were then calculated for each survey respondent, to be used in further analyses.

3.2.2 Model Fit

As Stares notes: "model selection is a notoriously difficult topic, with no consensus in the literature on the best way to approach it" (2008:96). As such, I will present a selection of the diagnostics provided by the software, alongside a newer test of absolute model fit, as recommended by Stares (2008) and Bartholomew et al. (2002), here termed the 'large marginals statistic' (LMS). As the substantive purpose of using latent trait modelling was to develop scales for various aspects of HIV knowledge, attitudes and practices, and to generate scores for individual surveys respondents on those scales, due attention was also paid to the parsimony and interpretability of the models produced, rather than having a monolithic focus on goodness of fit *per se*. The following is a brief summary of the statistical criteria used in model selection:

- L^2 – the likelihood ratio chi-squared statistic

- d.f. – the number of degrees of freedom

- BIC – the Bayesian Information Criterion = L^2 – log (N) d.f. (where N = number of cases)

- AIC – the Akaike's Information Criterion = L^2 – 2d.f.

- LMS – this test assesses absolute model fit, by calculating the percentage of standardised residuals in a model that are >4

L^2 is sensitive to large samples, and also does not follow a chi-squared distribution when large numbers of binary variables result in expected frequencies of response patterns in the models being less than 5, and so is problematic here, but reported alongside the other fit statistics. Both BIC and AIC are useful for large sample sizes and also favour more parsimonious models with fewer parameters set free to be estimated. A lower number indicates a better fit. LMS gives a percentage value for each model, again lower values meaning better fit.

3.2.3 Scale Development

Table 3.1 below shows the four key areas for scale development from the survey, further broken down into the more specific indices/scales that were estimated, and the survey items that contributed to those indices. Details of the individual survey items used in each scale may be found in Appendix 2. The decision to make several sub-scales for knowledge and attitudes was based on two key considerations. First, substantively, the large number of items involved precluded the development of unidimensional scales that were still meaningful – there was concern that subtleties would be lost from the data. Second, like factor analysis, latent trait analysis is a form of "data reduction" - presenting large amounts of data in a more condensed form. It was hypothesised that the fit of the traits would become very poor and their interpretation might become more difficult if all knowledge or attitude items were included in single scales. This was borne out in preliminary analyses, suggesting that the best approach would be to have a number of scales for different aspects of knowledge and attitudes.

Table 3.1: Scale names and survey items included

Key area	Scale Name	Description	Survey questions	Number of items in index
Knowledge	know_1	Index of knowledge of HIV in China in general	Q1, 8, 9, 11	4
	know_2	Index of knowledge of non-sexual routes of HIV transmission	Q401-5, 409-11	8

Key area	Scale Name	Description	Survey questions	Number of items in index
	know_3	Index of knowledge of sexual routes of HIV transmission	Q405-8, 412	4
	know_4	Index of knowledge of ways to avoid contracting/ transmitting HIV	Q701-8	8
Attitudes	Att_1	Index of attitudes to PLWHA	Q1901-7	7
	Att_2	Index of attitudes about HIV in China in general	Q20-24	8
Behaviours/ Practices	behav	Index of HIV-related behaviours	Q25-6, 31, 33-4	5
Media Exposure	media	Index of media exposure	Q3501-7	7

3.3: Further Analyses: Regression

The posterior scores for each survey respondent on each of the latent traits were then used as dependant variables in multiple linear regression analyses using the SPSS 16.02 software package (SPSS Inc., 2007) to explore whether being a migrant, while controlling for other variables in the model would have a significant effect on the levels of the latent traits estimated in the IRT modelling. All of the independent socio-economic variables were entered into the model as dummy variables, with the first level of each used as the reference variable (see table 4.1). As the aim of the analysis was to test the significance of the role of migration and other socio-economic variables, all were included in the models. In using regression analysis, one of the key advantages is that potential confounders are controlled for (providing the model is specified correctly). It would never be possible, nor necessarily desirable to control for everything that might have an effect on the dependant variable, but here given that the theory suggests that socio-economic factors would have a role in KAB/P and media levels (Chen et al., 2003), they were all retained. This form of analysis is frequently used to make inferences from a sample to the population, however, it must be noted that here, given a non-random sampling strategy was employed, any inferences are made very cautiously, and with the proviso that there are likely to be biases in the data that cannot be ignored.

Following the work of Rao et al. (2006), further multiple logistic regression analyses were performed using SPSS, using the eight scales a alongside the socio-economic variables as independent variables, with self-reported condom use ("ever used condoms") as the dependant variable. Again, all variables were retained in the model, as the substantive interest was to assess their different levels of significance with regards to condom use, while controlling for the other socio-economic variables in the model. The same cautionary note regarding inference equally applies here.

4. Results and Discussion

This section aims to provide a summary of the main findings of this study, set within the context of the research questions outlined in section 2.6, along with any implications that these findings may have. Both methodological and empirical issues will be discussed, recalling the twin aims of the study: namely to apply a novel technique to the analysis of KAB/P surveys, and then to utilise that technique to test some substantive questions regarding the relationship between migration and HIV risk in China.

4.1 Descriptive statistics

Table 4.1: Distribution of Socio-economic variables in survey sample

Variable	Categories	Total number	%
Income (RMB) per month	1 -1000	656	10.3
	1000-1999	1544	24.2
	2000-2999	1165	18.3
	3000-4999	674	10.6
	5000-7999	277	4.3
	8000-9999	44	0.7
	10000-14999	65	1.0
	15000-19999	8	0.1
	<20000	23	0.4
	Refuse	491	7.7
	No income	1435	22.5
	Total	6382	100.0
Education level	<elementary	252	3.9
	Junior high	1187	18.6
	High school	1750	27.4
	College	1184	18.6
	Bachelor degree	1855	29.1
	>master's degree	154	2.4
	Total	6382	100.0
Age	15-17	134	2.1
	18-24	2266	35.5
	25-49	3982	62.4
	Total	6382	100.0
Marital Status	Single	1726	27.0
	Unmarried with partner	1322	20.7
	Married	3280	51.4
	Divorced	29	0.6
	Widowed	15	0.2
	Total	6382	100.0
City	Beijing	1033	16.2
	Shanghai	1044	16.4
	Shenzhen	1086	17.0
	Wuhan	1067	16.7
	Zhengzhou	1064	16.7
	Kunming	1088	17.0
	Total	6382	100.0
Sex	Female	2990	46.9
	Male	3392	53.1
	Total	6382	100.0
Employment status	Youth	1604	25.1
	Blue Collar	1649	25.8
	White collar	1600	25.1

Migrant		1529			24.0		
Total		6382			100.0		

The survey involved a quota sample of individuals using telephone and face to face interviews (CHAMP/ UNAIDS, 2008), resulting in roughly even spread across employment types and city of residence. There was a slightly higher number of men in the sample, and the vast majority of the sample had at least completed junior high school. The average monthly income for the sample (taking the midpoint of each range) was approximately 1990 RMB. Only those between the ages of 15 and 49 were sampled; people outside this group are not considered to be at high risk of HIV infection. In the sample overall, approximately 57% reported ever using condoms.

4.2 Latent Trait Models

Table 4.2 shows the fit statistics for the chosen latent trait models used to create the indices for the different aspects of HIV KAB/P and media exposure, alongside the correlations between the individuals' scores on the latent traits and their sum scores for each set of items.

Table 4.2: Goodness of fit statistics for indices

Scale label	Scale name	L²	d.f.	BIC	AIC	LMS	Correlation[3]
Know_1	Index of knowledge of HIV in China in general	31.92	13	-81.98	5.92	0.00	0.656
Know_2	Index of knowledge of non-sexual routes of HIV transmission	543.73	238	-1541.45	67.73	4.12	0.960
Know_3	Index of knowledge of sexual routes of HIV transmission	130.55	20	-44.67	90.55	9.09	0.835
Know_4	Index of knowledge of ways to avoid contracting/ transmitting HIV	690.32	552	-4145.88	-413.68	4.27	0.837
Att_1	Index of attitudes to PLWHA	5270.18	2164	-13689.14	942.18	14.74	0.976
Att_2	Index of attitudes about HIV in China in general	712.89	481	-3501.27	-249.11	10.81	0.730
Behav	Index of HIV-related behaviours	565.14	92	-240.89	381.14	11.11	-
Media	Index of media exposure	812.83	113	-117.19	586.83	11.40	0.852

n.b. Behaviour index excluded from correlation analysis as the non-binary response options make sum-scoring

inappropriate. All correlations shown here were significant at the 1% level.

4.2.1 Goodness of Fit

Table 4.2 shows that the goodness of fit for the models varies considerably, with the knowledge traits generally performing better than the other models. As noted in section 3.2.2, lower scores of BIC and AIC both indicate better fit, here BIC is low for all scales. The LMS score, measuring the percentage of large standardised residuals in the model, and recommended as an alternative measure of fit (Stares, 2008), shows very good fit for know_1, 2 and 4, and good fit for all the remaining models except perhaps Att_1. As noted,

---

3   Pearson Correlation coefficient between the Index score and simple sum scores for each set of survey items. All correlations are significant at the 1% level. The correlation for 'Behav' is missing as the non-binary categorical answer options for these questions in the survey prevent the calculation of a meaningful sum score.

however, the substantive meanings of the traits is of greater relevance here, including the functioning of individual items, rather than the fit of the models overall.

The functioning of KAP/B items is an important consideration, as much value is placed on the predictive powers of these surveys in estimating people's levels of risky behaviours, and hence according to theory at least, their vulnerability to HIV infection. Furthermore, given that in this analysis a different method has been used to develop scales to the usual sum scoring, it would also be useful to draw some comparisons between the results of the two.

4.2.2 Research Question 1:
- How well do KAB/P survey items function in measuring what they are designed to measure, and to what extent is latent trait an improvement on CTT sum scores?

4.2.2.i Item Functioning
First, all of the correlations are high between sum and index scoring in the table. This might bring in to question the necessity of using IRT models when there are clear similarities between their results and those of the alternative more simple method. However, the best fitting model, Know_1, has the lowest correlation to the corresponding sum score, and the worst fitting, Att_1 has the highest correlation. These apparently anomalous results hint that there are hidden patterns in the data worthy of further investigation.

Figure 4.1 shows the item ICC's for Know_1 as an illustration to the discussion that follows. The ICC's for all of the remaining indices can be found in Appendix 3. The items were coded in such a way that the broken line in the ICC's, when questions relating to factual information, represents the correct response.

The ICC's for Know_1 (general knowledge about the HIV situation in China) shows that three out of the four items in the index are functioning correctly[4], with varying degrees of difficulty and discrimination. For example, the question: "Is HIV visible?" shows that for someone located towards the lower end of the knowledge latent trait, there is a very low probability of answering the question correctly, with the median individual (scoring 0 on the latent trait) having approximately 30% probability of giving the correct answer – this shows that this is quite a "difficult" item.

---

4   I.e. The items measured what they were supposed to – for example higher levels of correct knowledge answers for an individual would result in a higher score on the knowledge scale.
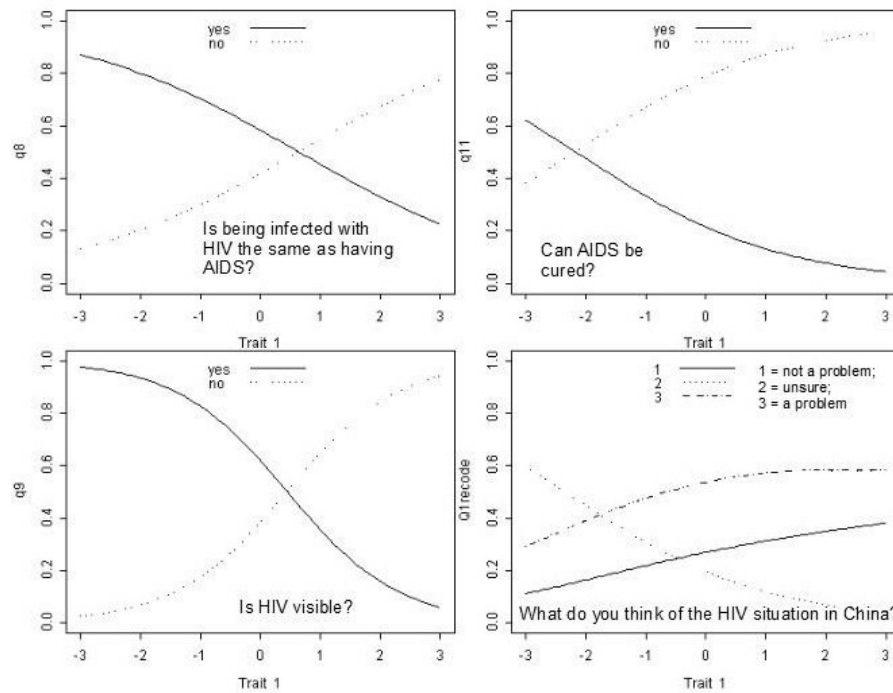
Figure 4.1: Item Characteristic Curves for Know_1

In order to be more likely to answer this question correctly than incorrectly, a person would only have score 0.5 higher on the trait than the median individual (with a zero-score). This demonstrates relatively high discrimination for this item (a steep slope). The question: "Is being infected with HIV the same as having AIDS" shows a similar level of difficulty to the previous question, but with a lower level of discrimination, indicating that this item does not do much to differentiate between people at with different levels of knowledge. The ICC for "Can AIDS be cured?" shows that the majority of people are likely to answer this question correctly – only those at the very lowest end of the latent trait being more likely to say "yes" - much lower levels of difficulty than for the previous two items. The ICC for the item regarding the HIV situation in China presents an interesting case. Here, as we move to higher levels of the latent trait for general knowledge about HIV, we can see that the probability of giving an "unsure" answer declines significantly, while the probabilities of giving either of the other two answers increases. This is perhaps counter-intuitive, although it would seem that this item regards strength of opinion rather than knowledge *per se*. What this tells us is that with increased knowledge of the HIV situation (the latent trait), comes greater certainty about the HIV situation in China. On reflection, this makes sense – with relatively low overall HIV prevalence in China, one might argue that someone who is knowledgeable about HIV might perceive the Chinese epidemic to be much less severe those of other countries. In essence, this question's subjectivity makes it difficult to interpret in the Chinese context, and would suggest that it might need to be rephrased in future Chinese KAB/P surveys. A possible alternative could be to ask: "What do you think the prevalence of HIV in China is?" with options ranging from 1 in 10 to 1 in 1 million.

From the ICC's it is evident that some items have greater difficulty and discrimination powers than others in the latent scales. Table 4.3 below presents a qualitative breakdown of each unusually performing ICC from the 7 remaining scales. A number of the items showed that they discriminated poorly between individuals on the traits meaning they could be considered for exclusion from future indices or be rephrased or replaced in future surveys on KAB/P in China (Schultz & Sibberns, 2004). Furthermore, by including only the best-functioning items, it might prove possible to develop shorter surveys that are still able to measure the constructs of interest to the requisite levels of accuracy, something that could help improve response rates.

Table 4.3: Unusual ICC's for the 7 remaining indices

Trait	Item	Item Description	ICC description
Know_1		See discussion above	
Know_2	401	Sharing needles	**Poor discrimination – flat ICC**
	404	Mosquito bites	Poor discrimination
Know_3	405	Kissing/ touching person with HIV	**Poor discrimination**
	412	Having unprotected sex with someone who looks healthy	Poor discrimination
Know_4	703	Can avoid HIV by not eating with infected person?	Poor discrimination
	707	Can avoid HIV by using a mosquito net	**Poor discrimination**
Att_1		All ICC's appeared to function well	
Att_2	2203	If one of your friends/relatives was infected with HIV, would you go with them to the clinic?	**Poor discrimination**
	2204	If one of your friends/relatives was infected with HIV, would you seek more information on HIV?	**Poor discrimination**
	2205	If one of your friends/relatives was infected with HIV, would you let them know you would care for them?	**Poor discrimination**
Behav	31	Have you had sex with someone other than your partner or spouse in the past 6 months?	Probability of answering "yes" increases slightly at higher end of trait
	33	If you get a new partner, would you use a condom the first time you have sex?	Low item difficulty
	34	Have you had multiple sexual partners in the last year?	**Poor discrimination**
Media	3501	How often do you have access to TV?	**Poor discrimination**
	3502	How often do you have access to Radio?	Poor discrimination

Note: descriptions in **bold** font represent strong effects.

4.2.2.ii Latent Trait versus Sum Scoring

The weightings estimated for each item offer another advantage of using IRT over CTT. These weights indicate that different questions have different levels of importance in measuring the trait. For example, in the second attitudes scale, the question: "If one of your friends/relatives were infected with HIV would you go with them to the local clinic?" Had a very low weight, indicating that it contributed very little to posterior scores of respondents on the scale. Using sum scoring, this item would have been "worth" the same amount as all of the other attitude variables in measuring attitudes to HIV in China, which might have oversimplified what the data was showing. It would certainly be useful to apply this analysis to other KAB/P surveys in

China in order to see whether issues to do with item functioning are the same across other samples.

A further advantage of the latent trait approach is that it permits estimation of item functioning for non-binary categorical item responses. This is something that sum scoring cannot do, as there is no clear numerical value that can be placed on the responses. This case is exemplified by the question: "if you get a new partner, would you use a condom the first time you had sex?", with answers "yes", "no" and "it depends" on the behaviour index - sum scoring might arbitrarily assign these categories values such as 1, -1 and 0, but the ICC's show that the second two answers are virtually indistinguishable from one another on the latent trait, and so should be scored accordingly.

The distributions of scores in the sample also demonstrate differences between sum and latent scores, many showing a departure from a normal distribution, once items are weighted. Graphs of the comparative distributions for each trait may be found in Appendix 4. The scales relating to sexual and non-sexual transmission knowledge, and knowledge of ways to avoid transmission all have a stronger negative skew than their corresponding sum scores. Both attitude and behaviour scales show a trimodal non-normal distribution using the latent trait scores, implying that there are three relatively distinct groups on these traits. The media scores again show a bimodal distribution, with two fairly distinct groups appearing. These distributions were concealed in sum scoring, and this might obscure particular groups with very low levels of KAB/P or media exposure that could be targeted by specific interventions.

One criticism is that latent trait scoring is less intuitive than the CTT alternative, and so it could be argued that a greater degree of the interpretation rests on the part of the researcher. This raises epistemological questions about the construction of knowledge that, while beyond the scope of this paper, nonetheless deserve mention. Acknowledging the inherent constructivist nature of modelling, arguably alongside that of all survey research (Bryman, 1984), I would argue that the results presented here should be seen within the context of all social science research – the scales cannot claim to be flawless reflections of the fluid concepts of knowledge and attitudes, but hopefully still permit us to learn something about the internal thoughts and behaviours of the people being studied.

4.3 Linear Regression

The 8 indices developed through latent trait analysis were used as dependant variables in a multiple linear regression, with the socio-economic characteristics of the respondents as independent variables. The results are presented in Table 4.4 below.

Table 4.4: Results of linear regression with indices as response variables

		Know_1	Know_2	Know_3	Know_4	Att_1	Att_2	Behav	Media
					Index name (dependant variable)				
Income (RMB) per month	1-1000 (ref)								
	1000-1999	0.048 **	0.042	0.047 *	0.048 *	0.083 **	0.099 ***	0.016	0.023 ***
	2000-2999	0.087 ***	0.048	0.012	0.064**	0.070 *	0.122 ***	0.037	0.103 ***
	3000-4999	0.074 ***	0.040	0.029	0.041	0.038	0.083 **	0.120 ***	0.094 **
	5000-7999	0.045	0.080	-0.034	0.015	0.069	0.120 **	0.156 ***	0.108
	8000-9999	-0.005	0.060	0.058	0.208*	0.168	0.069	0.256 **	-0.066
	10000-14999	0.062	0.064	-0.081	-0.037	0.283 **	0.072	0.221 **	0.000
	15000-19999	-0.029	-0.090	0.211	0.019	-0.499	-0.103	0.382	0.474 *
	<20000	0.064	-0.044	0.254 *	-0.204	0.042	-0.212	0.153	-0.051
	No income	0.168 ***	0.204 ***	0.097 ***	0.118 ***	0.248 ***	0.140 ***	0.033	0.004
Education level	<elementary (ref)								
	Junior high	0.038	0.111 **	0.073	-0.012	0.128 **	0.186 ***	0.059	0.226 ***
	High school	0.133 ***	0.275 ***	0.178 ***	0.020	0.218 ***	0.363 ***	0.221 ***	0.459 ***
	College	0.230 ***	0.395 ***	0.180 ***	0.108**	0.297 ***	0.480 ***	0.286 ***	0.669 ***
	Bachelor degree	0.251 ***	0.493 ***	0.295 ***	0.132 **	0.373 ***	0.541 ***	0.361 ***	0.706 **
	>master's degree	0.275 ***	0.480 ***	0.278 ***	0.072	0.362 ***	0.483 ***	0.395 ***	0.686 ***
Age	15-17 (ref)								
	18-24	0.014	0.121 *	0.280 ***	0.281 ***	-0.120	-0.054	0.011	0.092
	25-49	0.014	0.225 **	0.237 ***	0.260 ***	-0.115	0.018	-0.003	0.026
Marital Status	Single (ref)								
	Unmarried with partner	0.034 *	-0.020	0.077 ***	0.074 ***	-0.031	0.006	0.100 ***	0.080 ***
	married	0.025	-0.023	0.087 **	0.128 ***	-0.188 ***	-0.138 ***	-0.041	-0.058 *
	Divorced	-0.057	-0.198	0.077	-0.065	-0.274 *	-0.270 **	0.085	-0.101
	Widowed	0.010	-0.014	-0.010	0.199	-0.139	-0.011	0.335 *	0.125
City	Beijing (ref)								
	Shanghai	-0.043 *	0.051	-0.004	0.101 ***	-0.168 ***	-0.275 ***	-0.069 **	0.058 **
	Shenzhen	-0.157 ***	-0.162	-0.004	-0.238 ***	-0.247 ***	-0.255 ***	0.151 ***	-0.086 ***
	Wuhan	0.068 ***	-0.041	0.047	-0.097 ***	-0.279 ***	-0.363 ***	-0.041	-0.086 ***
	Zhengzhou	-0.070 ***	0.032	-0.147 ***	0.010	-0.163 ***	-0.172 ***	0.081 ***	-0.110 ***
	Kunming	-0.112 ***	0.157	-0.097 ***	-0.177 ***	0.120 ***	-0.042	0.343 ***	-0.047
Sex	Female (ref)								
	Male	0.066 ***	0.015	0.022	0.084 ***	0.091 ***	0.068 ***	0.102 ***	-0.020
Migrant status	Non-migrant (ref)								
	Migrant	0.100 ***	-0.291 ***	0.196 ***	-0.163 ***	-0.496 ***	-0.548 ***	0.123 ***	-0.326 ***
	R²	0.062	0.130	0.031	0.064	0.174	0.257	0.090	0.204

Significance levels: *** = 1%; ** = 5%; * = 10%

- To what extent do migrants' KAB/P and media exposure differ from those of non-migrants?

The results of the multiple linear regression show that in this sample, migrants scored significantly differently at the 1% level from the non-migrant group in all of the KAB/P and media indices. They show that migrants have higher levels of broad knowledge about HIV in China (know_1), higher levels of knowledge about sexual routes of HIV transmission (know_3) and significantly higher scores on the behaviour index. The results are somewhat mixed, however, as the group were also found to have lower levels of knowledge about non-sexual routes of transmission (know_2), lower levels of knowledge about ways to avoid contracting HIV (know_4), poorer attitudes on both scales, and lower levels of exposure to the media. The implications or these results are that in certain areas of knowledge, migrants appear to outperform their non-migrating counterparts, and that their HIV-avoiding behaviours appear to be greater than for the other group. This appears to contradict the theory that migrants tend to engage in more risky behaviours, outlined in section 2.4.4. However, it seems that migrants also have more negative attitudes to PLWHA and might be more likely to stigmatise this group. It has been suggested that higher levels of stigma can result in people being less likely to engage in health seeking behaviours (Lieber et al., 2006), meaning that people might be less likely to be tested for HIV, as well as motivating secretiveness and denial of risky behaviours, potentially allowing the epidemic to spread. It is questionable whether this conjecture is borne out by the these results, however, given that migrants appear to score more highly on the behaviour trait, despite poorer attitudes.

Migrants levels of media exposure were also significantly lower than the rest of the sample, and the implications of this are that this group may be more difficult to reach with health promotion messages, including campaigns to help fight stigma around sexual health and HIV. There is potentially then a problematic situation, in which migrants, with poorer attitudes to HIV, are also more difficult to reach in order to change those attitudes. These results also raise the issue of whether migrants' attitudes are more stigmatising because they have less access to mass media campaigns, or are there other factors at play? Qualitative research into the drivers of stigmatising attitudes might shed some light on this question and indeed an ethnographic study into Chinese HIV stigma describes it as a phenomenon that is firmly-rooted in notions of morality and "saving face" (Yang & Kleinman, 2008), which might mean that the any role the media has may come far behind these cultural factors.

Many of the other socio-economic variables in the model also show significant results for their role in

predicting levels of the KAB/P and media indices. A discussion of these results follows, and will form a contribution to the final research question: The results for income are mixed, broadly showing an expected increase in scores on all of the indices with increased income, in relation to the reference category. One of the possible reasons for non-significance in some of the middle income categories is the low number of cases for these in the sample (see Table 4.1). The "no income" results present a slight puzzle as they would be expected, if anything, to be lower than the reference category. Clearly there is something peculiar about this group that might warrant further investigation.

Having a higher level of education also broadly leads to a significant expected increase in scores on all variables, with higher levels of education generally being associated with higher scores. In addition, The effect of age appears to only be significant for the Know_2 to 4 indices, where being in a higher age group is associated with higher scores. For marital status, the results suggest that being in a relationship (either cohabiting or married) is associated with small positive effects on knowledge of sexual and non-sexual transmission routes of HIV (Know_2 and 3) compared with the reference category of "single". Furthermore, being married appears to be negatively associated with both attitude scales.

Geographic location appears to have a significant effect on people's scores on all scales besides Know_2. Broadly speaking, it seems that living in any of the five cities besides Beijing (the reference category) is associated with lower scores on know_1,2 and 3 (with the possible exception of Wuhan), and generally lower levels of positive attitudes (Att_1 and 2) and exposure to media. Despite this, respondents living in Shenzhen, Zhengzhou and Kunming all appear to be significantly more likely to score more highly on the behaviours index than their counterparts in the capital. The reasons for this result is unclear, but may be related to local HIV campaigns. On the issue of gender, being male is associated with small, but statistically significant increases in estimated scores on three of the four knowledge measures, both attitude measures and the behaviour scale compared to females. This might suggest that more female-focussed HIV campaigns could be fruitful, particularly given recent increases in the number of women infected in China (Xiao et al., 2006).

4.3.2 Criticisms of the Regression Models

Finally, a note about the R^2 values for the above eight regression models. It would seem from the results that none of these models explains a particularly large proportion of the variance in the dependant variables, and in fact the R^2 values for four of the models are very low (less than 10%). this presents a problem here as it demonstrates that a lot of the variation between individuals' scores on the KAB/P and media scales must be attributed to factors that have not been accounted for here. A conclusion though is that given many variables had highly significant results in the regression, the part of the variance in the dependant variables that the models do account for is explained well by the socio-economic factors, as suggested by the theoretical

framework (figure 2.1). I would argue that given the subjective nature of traits such as 'attitudes', 'knowledge', etc. we perhaps would never be able to explain their distributions in a sample fully through questions in a survey. However, this does call for more research into the determinants of KAB/P and media exposure that goes beyond socio-economic characteristics; an area that qualitative approaches might be very well suited to.

4.4 Binary Logistic Regression

Finally, binary logistic regression was performed to interrogate what effect, if any, the eight latent traits have on condom use, controlling for the socio-economic variables in the sample. This analysis attempts to combine the effects of all of the latent scores into one model, acknowledging that HIV-protective behaviours, such as condom use, are not proscribed by any single characteristic, but rather by combinations of traits, alongside individual behavioural mechanisms, and possibly some socio-economic factors. As all of the socio-economic categories had at least one significant odds ratio, all were retained in the model. The results of the analysis are presented below in table 4.5.

Table 4.5: Results of binary logistic regression with "ever used condoms" as dependant variable.

Variable		Odds Ratio	Significance (p value)
Indices	Know_1	1.141	0.037
	Know_2	1.100	0.036
	Know_3	1.112	0.019
	Know_4	1.359	<0.005
	Att_1	0.965	0.402
	Att_2	1.074	0.176
	Behav	1.381	<0.005
	Media	1.091	0.067
Income (RMB) per month	1-1000 (ref)		
	1000-1999	1.578	<0.005
	2000-2999	1.943	<0.005
	3000-4999	2.503	<0.005
	5000-7999	2.526	<0.005
	8000-9999	2.428	0.070
	10000-14999	12.352	<0.005
	15000-19999	0.960	0.964
	<20000	3.853	0.043
	No income	0.537	<0.005
Education	<elementary (ref)		
	Junior high	1.540	0.006
	High school	1.809	<0.005
	College	2.078	<0.005
	Bachelor degree	1.792	0.002
	>master's degree	2.620	0.001

Variable		Odds Ratio	Significance (p value)
Age	15-17 (ref)		
	18-24	3.508	0.018
	25-49	4.593	0.005
Marital Status	Single (ref)		
	Unmarried with partner	6.828	<0.005
	married	15.468	<0.005
	Divorced	13.781	<0.005
	Widowed	10.543	<0.005
City	Beijing (ref)		
	Shanghai	0.820	0.085
	Shenzhen	0.718	0.005
	Wuhan	1.091	0.452
	Zhengzhou	0.917	0.456
	Kunming	0.981	0.868
Sex	Female (ref)		
	Male	2.191	<0.005
Migrant Status	Non-migrant (ref)		
	Migrant	0.954	0.640

4.4.1 Research Question 3:

- What effect do the KAB/P and media exposure levels in the sample have on individual likelihood of using condoms?

These results show that all four of the knowledge scales, and the behaviour scale are significant predictors of condom use in the sample at the 5% level, and media exposure is significant at the 10% level, controlling for the other variables, Figure 4.2 shows the odds ratios and 95% confidence intervals for each of the eight indices, those that intersect the red line are not significant at the 5% level.
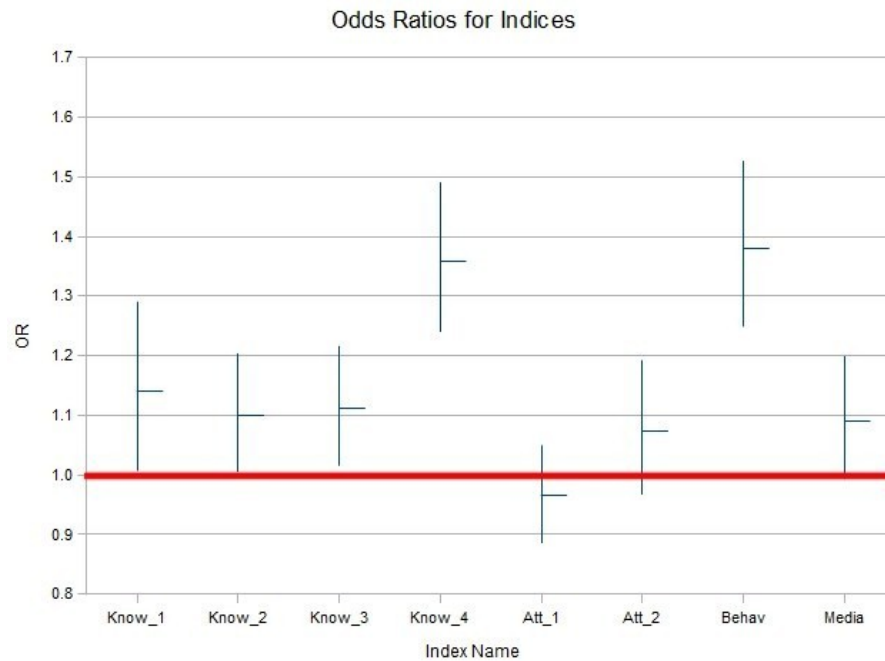
Figure 4.2: Odds ratios and 95% confidence intervals for the eight indices as independent variables in binary logistic regression.

The biggest effect estimated by the model is for the "knowledge of ways to avoid HIV transmission" scale (Know_4), a one unit increase of which is expected to result in a 35.9% increase in the odds of condom use, controlling for other variables. Interestingly, neither attitude index had a significant effect on the odds of ever using condoms, controlling for the other variables in the model. From the odds ratios, we can see that having higher levels of any of the significant latent traits is associated with a higher odds of ever having used a condom. Specifically, that a one unit increase in know_1, 2 and 3 are estimated to increase the odds by 14.1%, 10% and 11.2% respectively, and behav and media by 38.1% and 9.1% respectively. Given that the survey does not include data on time order of KAB/P, media or condom use, however, we cannot comment on causal links between these variables. One might posit that use of condoms could well be a pre-date knowledge of HIV or indeed be unrelated. If anything, we might argue that the commonly used indicator of condom use (cf. Sheeran & Taylor, 1999) be replaced with, or enhanced by other variables such as "ever visited CSW's" or "number of sexual partners".

Many of the other socio-economic variables in the model continued to be significant, indicating that these demographic factors continue to play a role in safer sex behaviour directly beyond that hypothesised through their effects on KAB/P and media exposure. Of particular note, however is that migrant status is not a significant predictor of condom use in this analysis, indicating that any differences between migrants and their non-migrating counterparts looks to have been accounted-for by the other variables, if it exists at all. This topic will be revisited in answering the final research question below.

Finally, the R^2 of this model was estimated to be 0.362, which while far from providing a perfect fit,

indicates that a fair proportion of the variance in self-reported condom use can be attributed to the influence of the various independent variables in the model. This could perhaps be expected, given the likely role that other factors, such as previous number of sexual partners, availability of condoms, etc., might have in creating 'noise' in the data. Future KAB/P surveys should consider the inclusion of these variables.

4.5 Research Question 4:

- Does the theoretical framework hold with regards to migrants? Are there intermediate variables that play a role in their HIV risk, such as socio-economic variables, or are migrants inherently psychologically different to non-migrants?

Drawing together threads from the previous two research questions, we will try to relate the substantive findings of this study to the theoretical positions outlined in section 2.5. Becker's (1977) health beliefs model underlines the importance of demographic and social factors in determining beliefs and hence behaviours, and this theory appears to be supported by this study. Socio-economic variables were largely found to have a significant role in predicting levels of KAB/P.

Regarding migrants specifically, a theoretical relationship has been posited between migrants and lower levels of KAB/P and hence higher propensity to undertake risky behaviours, alongside the theory that migrants are simply more risk-taking as individuals. The results from this study, however are mixed. While migrants do appear to have significant differences in their levels of KAP/P and media exposure, they do not score lower than non-migrants on all of these scales, as the theory would suggest, and in fact scored higher levels than non-migrants on the behaviour scale. Moreover, the predicted link between attitudes and behaviours (see Ajzen & Fishbein, 1977) also does not appear to hold in this sample of migrants, indicating perhaps that there are other factors, perhaps intermediating between attitudes and behaviours, as Fishbein (2000) postulates.

Controlling for levels of KAB/P, media and socio-economic variables allowed us to see whether there was some other underlying psychological characteristic that migrants possess that might make them more risk-taking, as Brockerhof and Biddlecom (1999) suggest. This again does not seem to hold for this group, given that their odds of condom use were no different from non-migrants. What is apparent from these results, though is that many of the socio-economic variables do continue to have a significant role as predictors of condom use, after controlling for the KAB/P and media

indices.

4.6 Criticisms

This survey was focussed on six cities in China, and as such the data must be considered within that context – it seems likely both intuitively, and from some of the results presented above, that geographical location plays a role in determining certain characteristics here, and as such more research is needed into KAB/P and media exposure using a country-wide sample. Second, the quota sampling strategy used here does is not a statistically random sample, and as such there may be biases in the data that cannot be estimated. This means that any of the conclusions drawn must be seen as relating to this sample, with any wider inferences cautioned against. It would certainly be interesting, however to extend the methodologies used here to a bigger and more representative sample to see if the results were replicated. This applies equally to the discussion of item performance in section 4.3 as it does to the results in later parts of section 4.

As noted in section 3.2.3, the selection of items for each of the scales was largely informed by theory, and in part by the necessity of having easily interpretable scales. This may, to the reader, appear rather arbitrary, and other combinations of items could equally have been chosen for the scales. Whether or not this would have made a significant difference to the results is open to debate, and future research could certainly focus on doing a comparative study of item functioning and scale effects, using different combinations of survey items in scale construction. A further extension could also be to use multidimensional scaling, rather than the simpler unidimensional scales here, to estimate interrelations between the latent traits.

And finally, it must be acknowledged that in trying to avoid taking the reductionist approach to migration cautioned against by Yang (2006), this study has inevitably been reductionist in many other areas, not least the highly important topic of gender in relation to HIV risk and social marginalisation. A better, and larger study would certainly focus the changing dynamics of gender in contemporary China, as it is in this area, as much as in the study of correlates of population movement that more work needs to be done to understand the drivers of the HIV epidemic.

5 Conclusions

This study set out to investigate a number of questions, both methodological and substantive. It has found that by applying latent trait analysis to the development of indices for HIV knowledge, attitudes, behaviours and media exposure permits large numbers of survey questions to be condensed in to a small number of scales, and that each survey question's functioning on that scale can then be individually assessed. It found that a number of the questions do not necessarily interrogate what we might have expected them to. A notable example regarded subjective perceptions of the severity of the HIV epidemic in China.

The study has also found that, in line with previous research, migrants do appear to have poorer levels of some areas of knowledge, attitude and media exposure compared with non-migrants. However, contrary to previous theory, the migrants in this sample appeared to have better knowledge about sexual routes of HIV transmission, and better levels of health-protective behaviours. This suggests that while migrants might be more likely to stigmatise HIV-positive people, this does not necessarily appear to be connected with how they report to behave in relation to avoiding HIV. Many mass media campaigns aim to reduce HIV stigmatisation, these results suggest that there may be two things happening here – first, that because of their generally lower levels of media exposure, migrants' attitudes to HIV remain more stigmatising than those of their non-migrating counterparts, and second, that migrants may be a more difficult group to reach through mass media campaigns in the future. The policy implications of this are that this group might need to be targeted in a different way in future in order to promote HIV-related information and attitudinal change, perhaps through localised campaigns in urban areas that house large numbers of migrants.

Finally, when using condom use as an outcome variable, the analysis suggested that controlling for levels of KAB/P, media exposure and socio-economic factors, migrants did not differ significantly from non-migrants. This does not mean necessarily that migrant's use of condoms is equal to that of non-migrants, but rather that any differences are due, in the models used here, to the other factors. This would imply that these migrants do not appear to have an intrinsic difference in their predisposition to epidemiologically risky behaviours (for example, the psychological traits suggested by Brockerhof & Biddlecom, (1999)), other than those that were accounted for in the KAB/P scales. In terms of interventions, therefore, I would suggest that attention is paid to increasing levels of KAB/P and media exposure for migrants. By assuming that migrants have some other psychological trait that makes them behave more riskily not only stigmatises them, but also might deflect attention from making interventions that could be very effective in reducing their risk of HIV infection, although any interventions must, in turn, be sensitive to avoiding further stigmatisation of this group.

Much work on the relative risks of HIV infection among migrants in China remains theoretical. The predicted explosion in new cases has by-and-large not yet materialised. This is likely to have been due to many factors, not least the plurality of structural and cultural differences between Chinese society and higher prevalence areas of the world. However, the adage that prevention is better than cure remains particularly significant while ever a cure remains elusive. In a country of over a billion people, undergoing rapid social and economic change, complacency could be disastrous. Targeting specific populations that are deemed to be at risk is essential, but more needs to be known about what makes them at risk in the first place if that targeting is to be a success.

Appendix 1: Latent trait model for non-binary categorical response variables

If observed variable i has 2 categories (I.e. is binary), then the model is defined as:

$$logit\ \pi_i(x) = [\log_e \pi_i(x)] / [1-\pi_i(x)] = \alpha_{i0} + \sum_{j=1}^{q} \alpha_{ij}x_j$$

(1)

Where: $\pi_i(x)$ is the probability of a particular response for item i, given a value on the latent variable; α_{i0} is the intercept, and α_{ij} is the loading for item i on latent variable j.

If observed variable i has $c_i > 2$ categories (s), where s takes the values $0,1,2,... c_i-1$, with 0 as the reference category, the model is defined as:

$$log\ [p_{is}(x)/p_{i0}(x)] = \alpha_{i0(s)} + \sum_{j=1}^{q} \alpha_{ij(s)} x_j$$

(2)

where $p_{is}(x)= P(y_i = s|x)$ is the probability of responding in category $s(= 1,2,... c_i-1)$ to item i, given a value or set of values on the latent trait, $p_{i0}(x)$ is the probability of responding in the reference category to that item, and $x =$ the vector of values for the q latent variables (taken from Stares, 2008).

In this study, the majority of items had binary response categories, and so the model specified by (1) was used. However, for the Att_1 scale, there were >2 potential response options, meaning that (2) was specified as the appropriate model.

Appendix 2: Survey questions used in scale development

Know_1 (General knowledge about the HIV situation in China)
- What do you think of the HIV situation in China? (problem/ unsure/ not a problem)
- Is being infected with HIV the same as having AIDS? (yes/no)
- Are there any visible symptoms of HIV infection? (yes/no)
- Can AIDS be cured? (yes/no)

Know_2 (Knowledge of non-sexual routes of transmission)
- Sharing needles with others (yes/no)
- Receiving a blood transfusion of unscreened blood (yes/no)
- Using the same toilets as someone who is HIV positive (yes/no)
- A mosquito bite (yes/no)
- Kissing on the cheek/ touching someone who is HIV positive (yes/no)
- An HIV positive person coughing or sneezing nearby (yes/no)
- Using the same drinking glass, chop sticks or eating together at the same table as a HIV positive person (yes/no)
- Working in the same room as a HIV positive person (yes/no)

Know_3 (knowledge of sexual routes of transmission)
- Kissing on the cheek /Touching someone who is HIV positive (yes/no)
- Having unprotected penetrative anal sex with a person who is infected with HIV (yes/no)
- Having unprotected penetrative vaginal sex with a person who is infected with HIV (yes/no)
- Having unprotected penetrative oral sex with a person who is infected with HIV (yes/no)
- Having unprotected sex with someone who looks healthy (yes/no)

Know_4 (knowledge of ways to avoid HIV infection)
- By having sex only with one faithful sexual partner who is not infected with HIV (yes/no)
- By abstaining from sex (yes/no)
- By not eating with infected person (yes/no)
- By using condoms when having sex (yes/no)
- By abstaining from penetrative sex (yes/no)
- By not using injection needles already used by others (yes/no)
- By using mosquito net (yes/no)
- By using only screened clean blood during a transfusion (yes/no)

Attit_1 (attitudes to PLWHA)
- Would you be willing to work in the same place with a person who is HIV positive? (willing/ unsure/ unwilling)
- Would you be willing to live in the same room/apartment as a person who is HIV positive? (willing/ unsure/ unwilling)
- Would you be willing to eat food at the same place as a person who is HIV positive? (willing/ unsure/ unwilling)
- Would you be willing to share tools or office equipment with a person who is HIV positive? (willing/ unsure/ unwilling)
- Would you be willing to shake hands with someone who is HIV positive? (willing/ unsure/ unwilling)
- Would you be willing to receive services from someone who is HIV positive, for example a barber, hairdresser, beautician? (willing/ unsure/ unwilling)
- Will you be willing to care for a relative sick with AIDS? (willing/ unsure/ unwilling)

Attit_2 (Attitudes to the HIV situation in China generally)
- Do you feel that HIV positive students/children should be allowed to study in school/university with uninfected children/students? (yes/no)
- People who have HIV/AIDS through sex or drug use have gotten what they deserve? (yes/no)
- If one of your friends/relatives was infected with HIV would you stop talking to them? (yes/no)

- If one of your friends/relatives was infected with HIV would you not touch them? (yes/no)
- If one of your friends/relatives was infected with HIV would you go with them to the local clinic? (yes/no)
- If one of your friends/relatives was infected with HIV would you seek more information on HIV? (yes/no)
- If one of your friends/relatives was infected with HIV would you ensure that they know that you would care for them if they are sick? (yes/no)
- Do you believe that a person with HIV can have a productive life? (yes/no)

Behav (Index of HIV-related behaviours)
- Have you sought information on HIV/AIDS voluntarily? (ever/ never)
- Have you discussed HIV/AIDS related issue with your family or friends or colleagues? (ever/ never)
- In the past 6 months, have you had sex with a person who is not your spouse or steady, single partner? (yes/ no)
- If you get a new partner, would you use a condom the first time you have sex? (yes/ no/ it depends)
- Have you had multiple sexual partners in the past year? (yes/ no)

Media (Index of exposure to different types of media)
- How often do you get access to these Media?
    - TV (at least weekly/ less than weekly)
    - Radio (at least weekly/ less than weekly)
    - Magazine (at least weekly/ less than weekly)
    - Newspaper (at least weekly/ less than weekly)
    - Internet (at least weekly/ less than weekly)
    - Cinema (at least weekly/ less than weekly)
    - Public monitor TV screen (at least weekly/ less than weekly)
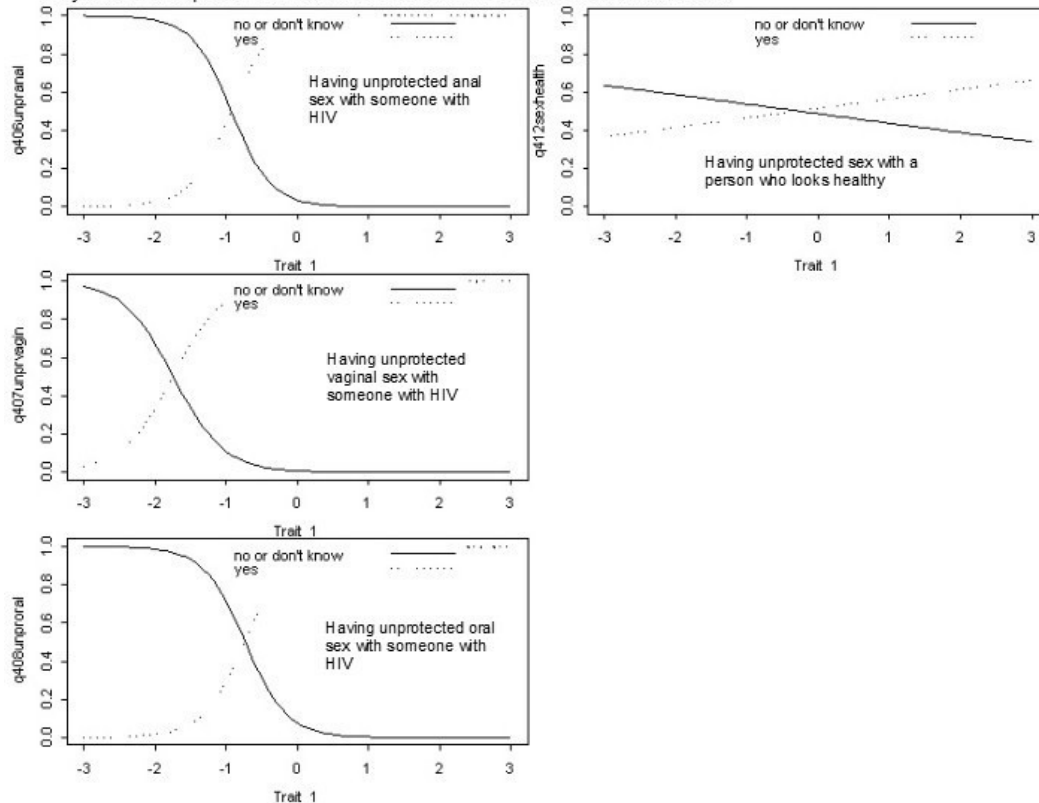
Appendix 3: ICC's for each latent scale

Know_2 – Knowledge of non-sexual routes of transmission

Do you believe a person can become infected under the circumstances below?
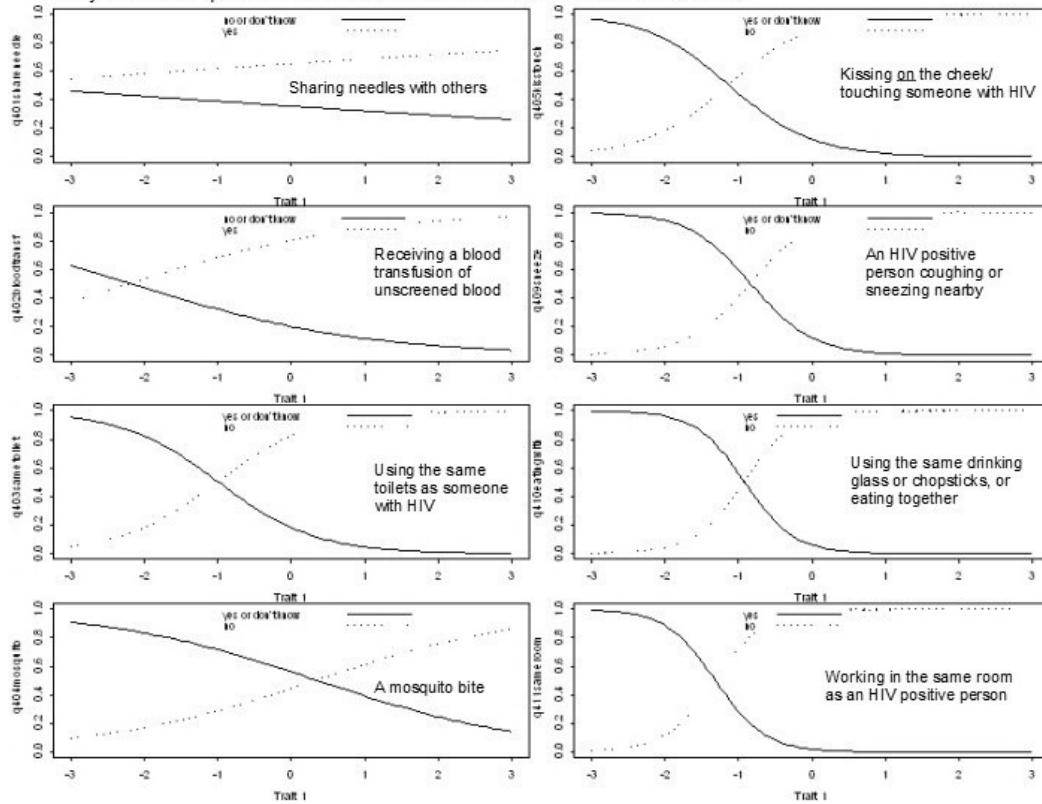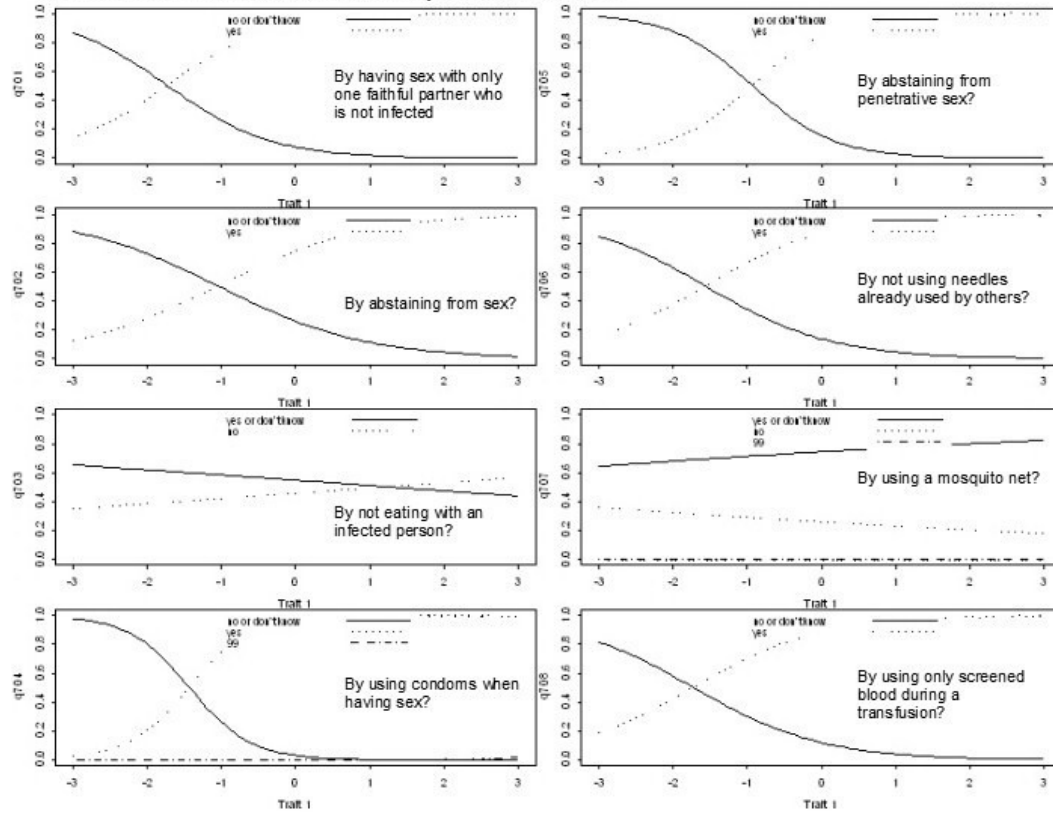

Know_3 – Knowledge of sexual routes of transmission

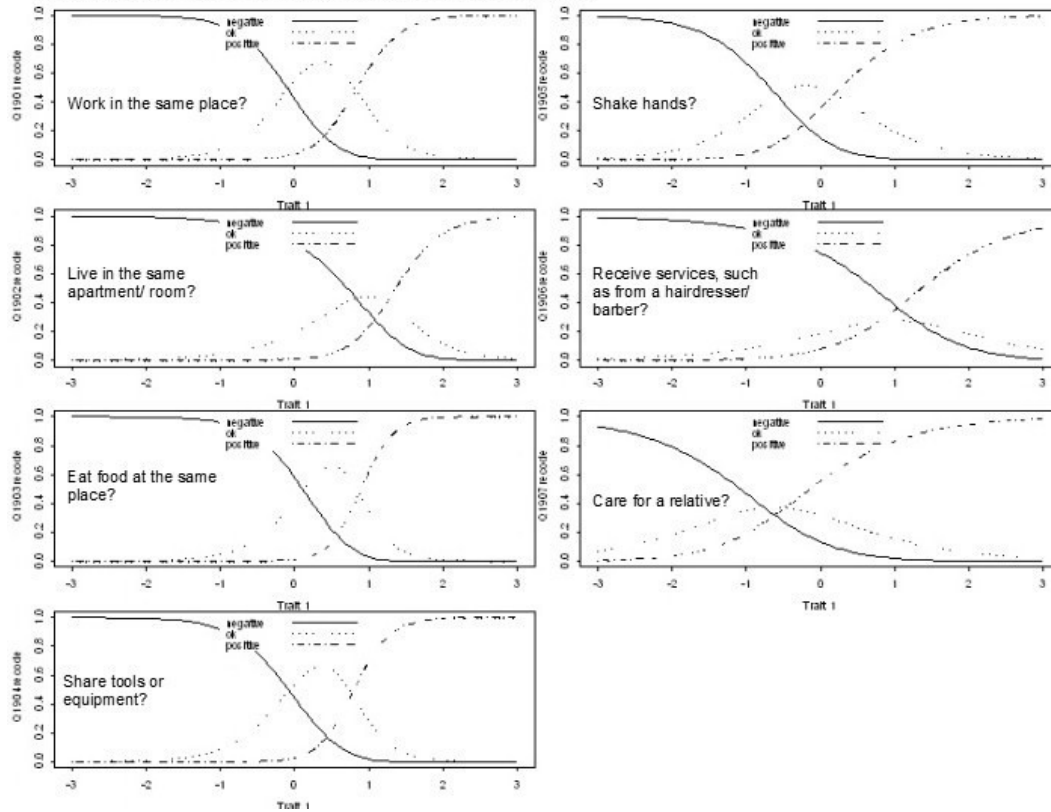Do you believe a person can become infected under the circumstances below?

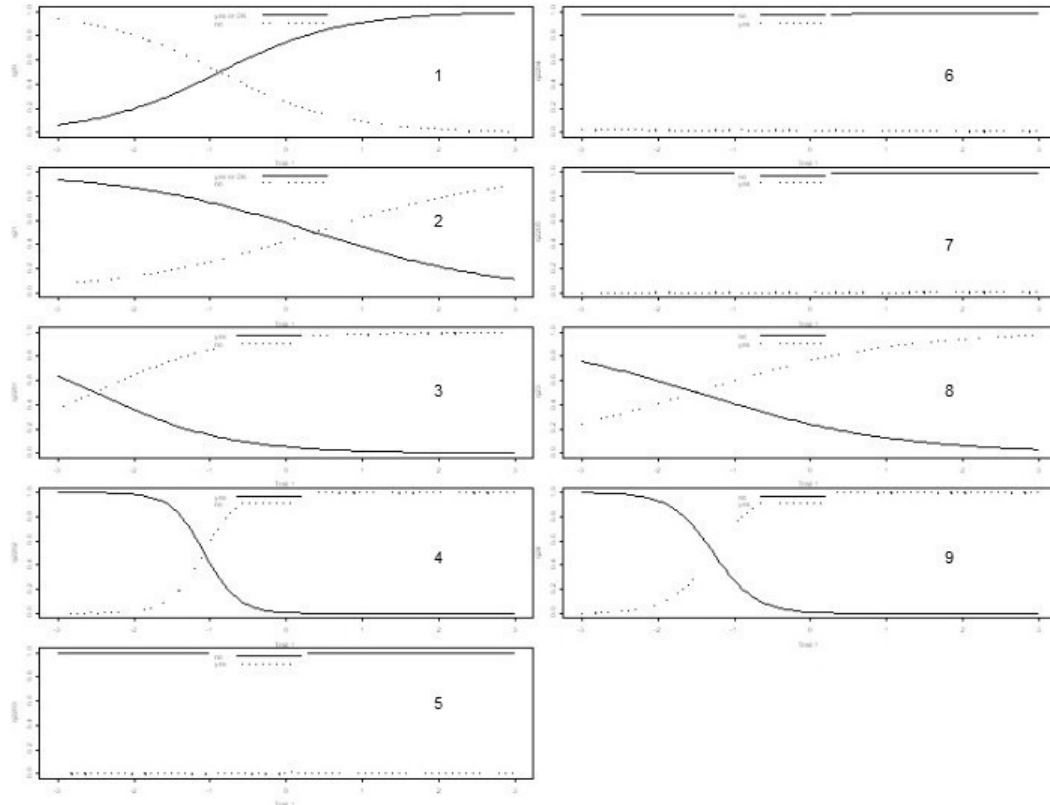Can the risks of HIV infection be reduced by the methods below?

Att_1 – Attitudes to people living with HIV/AIDS

Would you be willing to do the following with a person with HIV?

Att_2 – Attitudes to HIV in China

1 = Do you feel that HIV positive students/children should be allowed to study in school/university with uninfected children/students?
2= People who have HIV/AIDS through sex or drug use have gotten what they deserve?
3= If one of your friends/relatives was infected with HIV would you stop talking to them?
4= If one of your friends/relatives was infected with HIV would you not touch them?
5= If one of your friends/relatives was infected with HIV would you go with them to the local clinic?
6= If one of your friends/relatives was infected with HIV would you seek more information on HIV?
7= If one of your friends/relatives was infected with HIV would you ensure that they know that you would care for them if they are sick?
8= do you know where you can find out about HIV?
9= Do you believe that a person with HIV can have a productive life? (yes/no)

Behav – Index of HIV-related behaviours
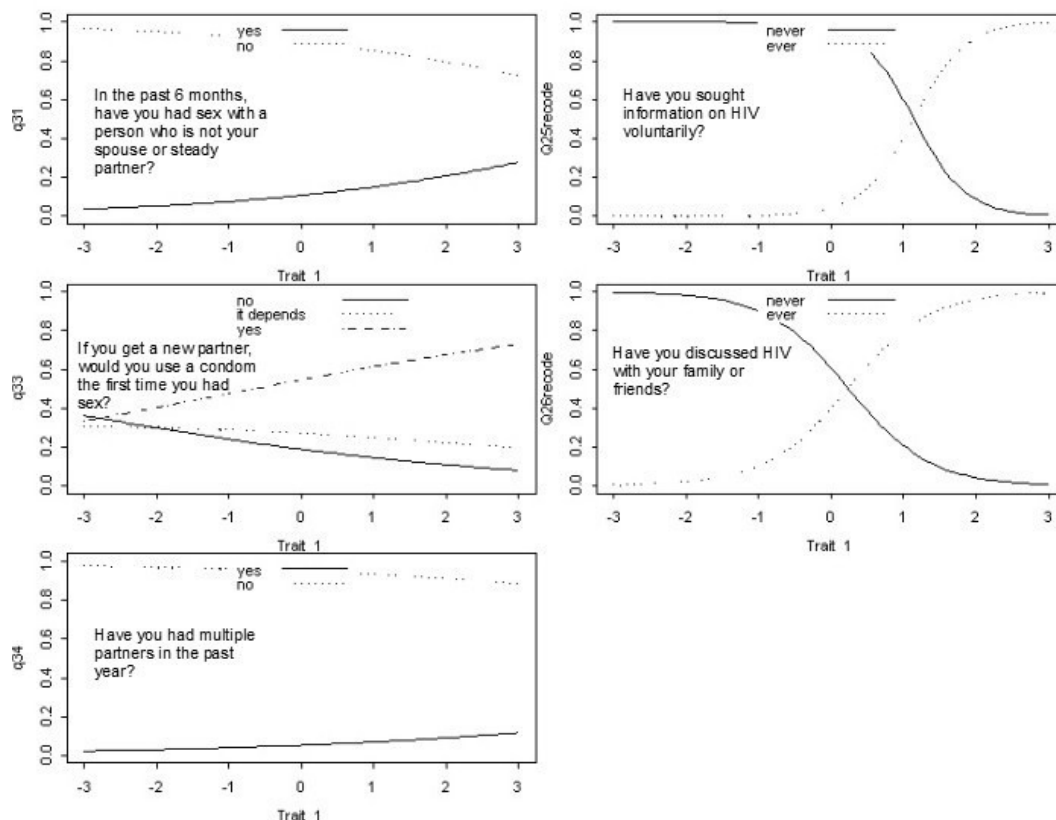
Media – Index of media exposure

How often do you get to these medias?

47
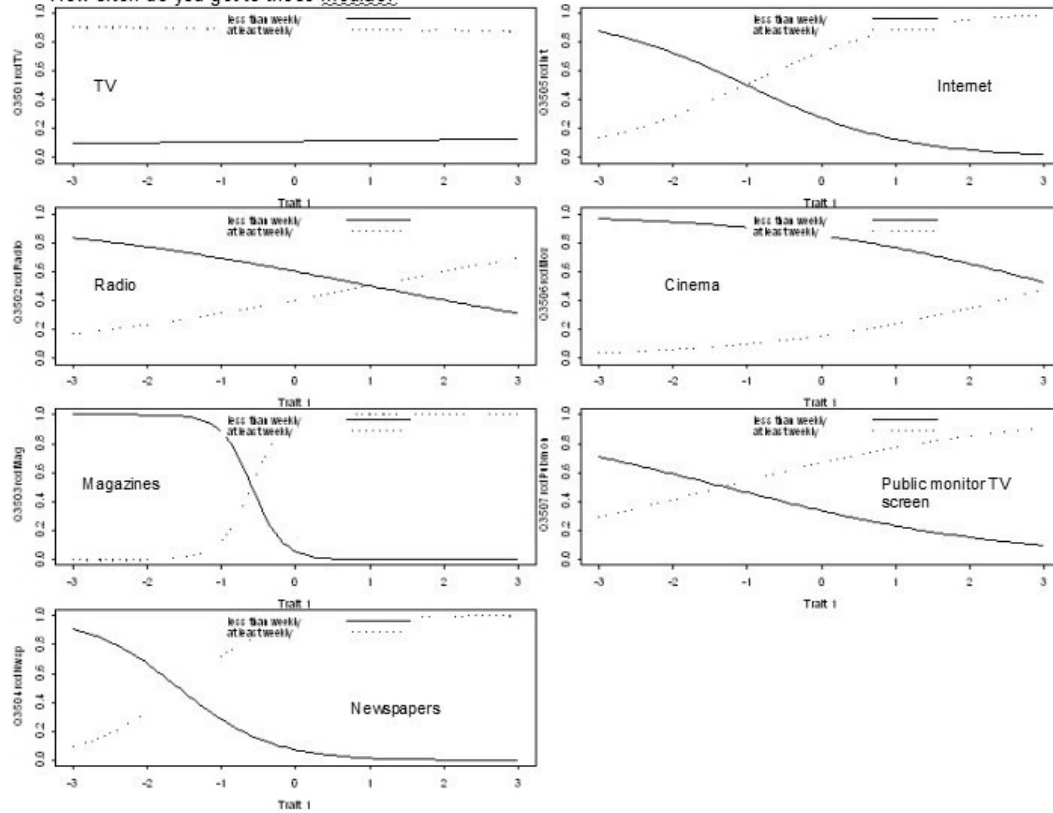
Appendix 4: histograms showing distributions of sum versus latent scores for 7 indices

Figure 1: Know_ and 2
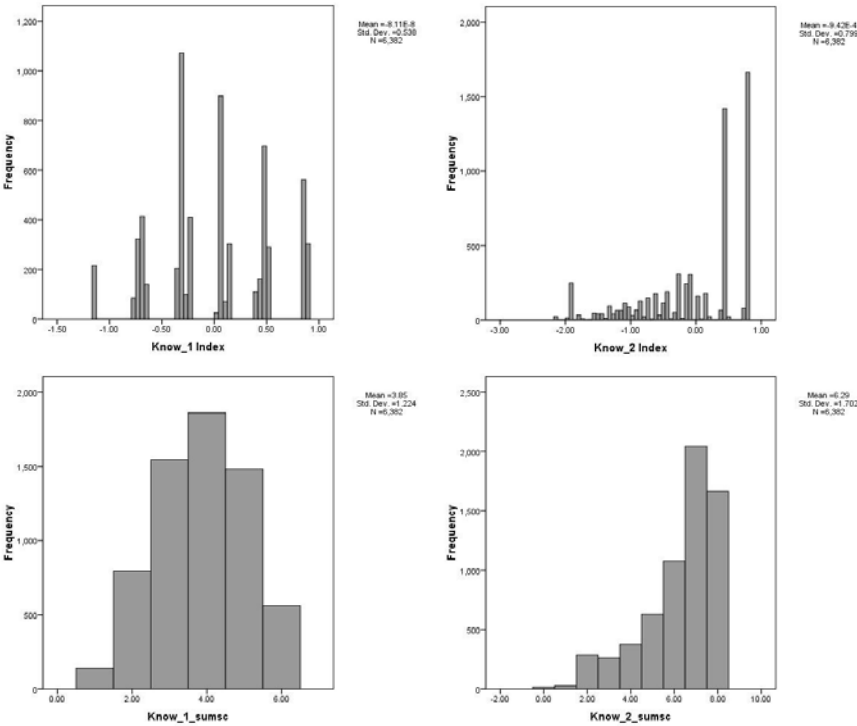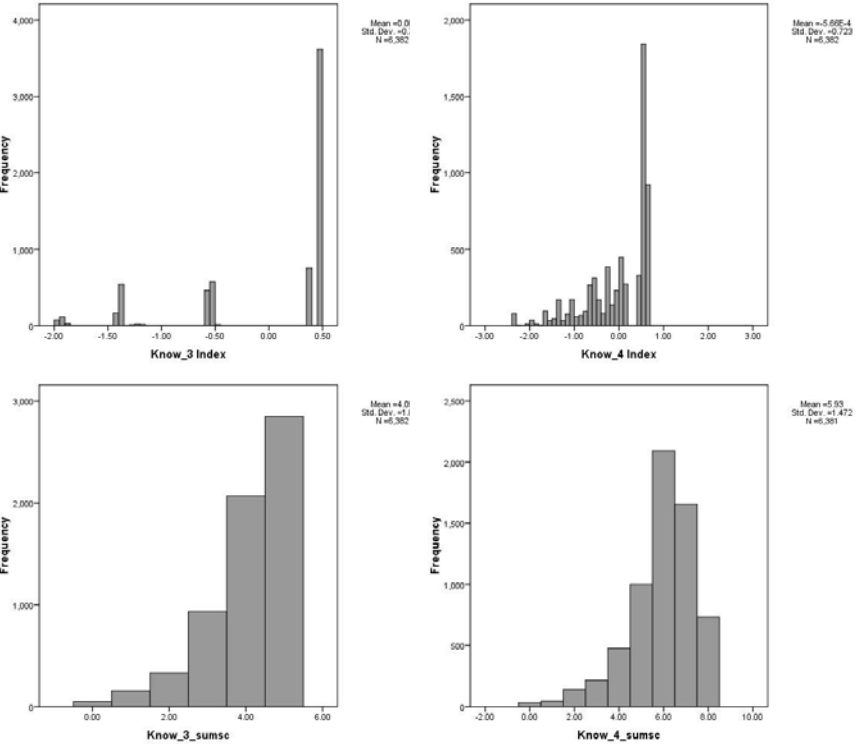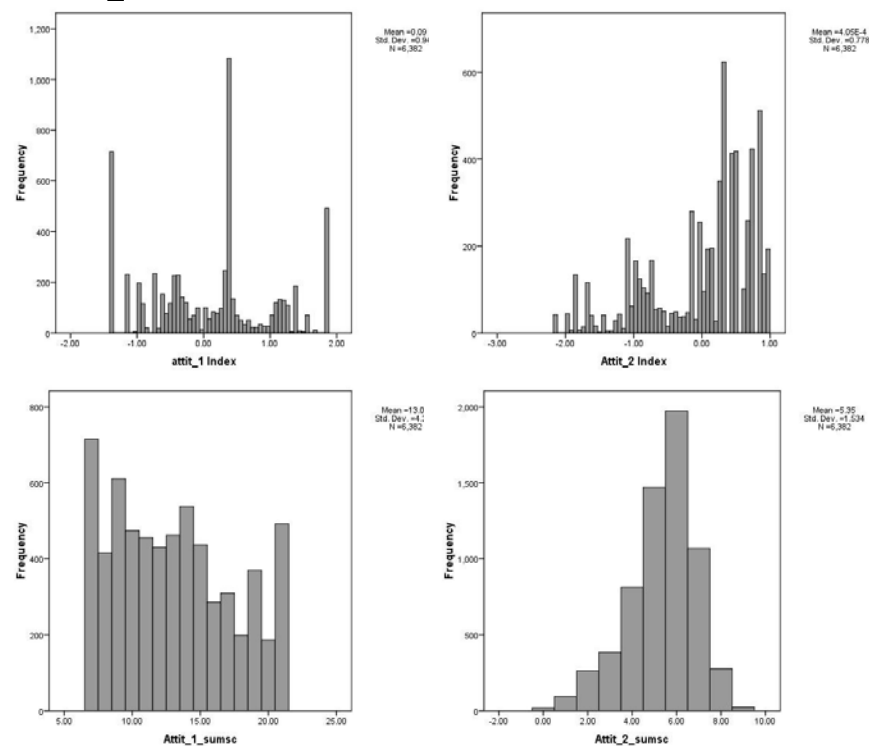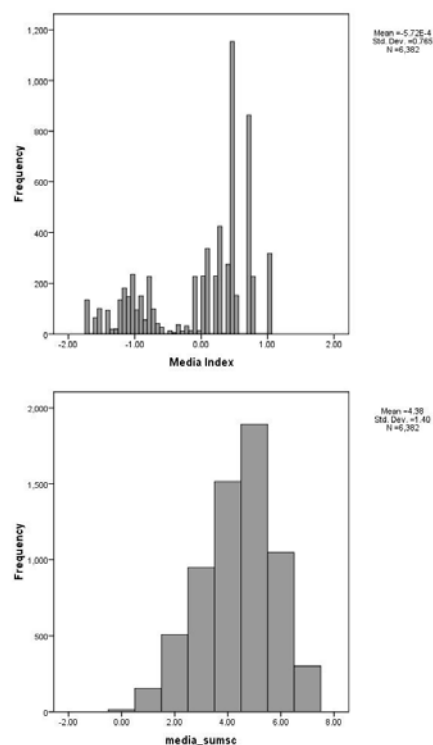

Figure 2: Know_3 and 4

Figure 3: Att_1 and 2

Figure 4: Media

Bibliography

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: a theoretical analysis and review of empirical researchers. *Psychological Bulletin, 84*(5), 888.

Bakken, B. (2000). *The exemplary society: human improvement, social control, and the dangers of modernity.* London: Oxford University Press.

Barnett, T., & Whiteside (Book). (2006). *HIV/AIDS in the Twenty -First Century* (2nd Edition). Basingstoke: UK: Palgrave Macmillan.

Bartholomew, D., & Knott, M. (1999). *Latent variable models and factor analysis*. London: Arnold.

Bartholomew, D., Steele, F., Moustaki, I. et al. (2008). *Analysis of multivariate social science data. 2nd Edition*. London: CRC Press.

Bartholomew, D., Steele, F., Moustaki, et al. (2002). *The analysis and interpretation of multivariate data for social scientists*. London: UK: Chapman and Hall.

Becker, M., Haefner, D., Kasl, S. et al. (1977). Selected psychosocial models and correlates of individual health-related behaviors. *Medical Care, 15*(5 (suppl.), 27-46.

Berger, B., Ferrans, C., & Lashley, F. (2001). Measuring stigma in people with HIV: psychometric assessment of the HIV stigma scale. *Research in Nursing and Health, 24*(6), 518.

Borowitz, M., Wiley, E., Saadah, F. et al. (2004). *Addressing HIV/AIDS in East Asia and the Pacific*. Washington, D.C: World Bank Publications.

Brockerhof, M., & Biddlecom, A. (1999). Migration, sexual behavior and the risk of HIV in Kenya. *International Migration Review, 4,* 833-856.

Bryman, A. (1984). The dabate about quantitative and qualitative research: a quesion of method or epistemology. *The British Journal of Sociology, 35*(1), 75-92.

CHAMP/ UNAIDS. (2008). AIDS-related knowledge, attitudes, behavior and practices: A survey of 6 Chinese cities. Beijing: China: CHAMP/UNAIDS.

Chan, K. (Presentation). (January 2008). Internal Labor Migration in China: trands, geographical distribution and policies. In Source: http://www.un.org/esa/population/meetings/EGM_PopDist/Chan.pdf.

Chen, J., Shengli, C., & Choe, M. (2003). Who has correct knowledge about HIV/AIDS in China?. *Asia-Pacific Population Journal, 18*(4), 25-49.

Darke, S., Hall, W., Heather, N. et al. (1991). The reliability and validity of a scale to measure HIV risk-taking behaviour among intravenous drug users. *AIDS, 5,* 181.

Dutton, M. (1998). *Streetlife China*. Cambridge, UK: Cambridge University Press.

Fendrich, M., Smith, E., Pollack, L. et al. (2008). Measuring sexual risk for HIV: A Rasch scaling approach. *Arch. Sex. Behav., online first.*

Fishbein, M. (2000). The role of theory in HIV prevention. *AIDS care, 12*(3), 273.

Garten, R., Lai, S., Liu, W. et al. (2004). Trends in injection drug use and the HIV epidemic in Southern China. In *AIDS in Asia* (379). New York: Kluwer Academic Press.

Gerbert, B., Bronstone, A., McPhee, S. et al. (1998). Development and testing of an HIV-risk screening instrument for use in health care settings. *American Journal of Preventative Medicine, 15*(2), 103.

Gillespie, S., Kadiyala, S., & Greener, R. (2007). Is poverty of wealth driving HIV transmission. *AIDS, 21*(S7), S5-16.

Gorbach, P., Ryan, C., Saphonn, V. et al. (2002). The impact of social, economic and political forces on emerging HIV epidemics. *AIDS, 16*(Suppl 4), S35-43.

Graf, M., Shao, Y., Zhao, Q. et al. (1998). Cloning and characterization of a virtually full-length HIV type 1 genome from a subtype B'-Thai strain representing the most prevalent B-clade isolate in China. *Chinese Reserch in Human Retroviruses, 14*(3), 285-288.

Grusky, O., Liu, H., & Johnson, M. (2002). HIV/AIDS in China:1990-2001. *AIDS and Behavior, 6*(4), 381-394.

Guomei, X. (2005). *HIV/AIDS in China*. Beijing, China: Foreign Languages Press.

Haitao, Y., Jian-Ping, D., & Yaping, D. (2004). Prevention of sexual STI/HIV transmission in Jiangsu Province, China. In *AIDS in Asia* (471). New York: Kluwer Academic Press.

Hennessey, M., Bleakey, A., Fishbein, M. et al. (2008). Validating an index of adolescent sexual behavior using psychological theory and social trait correlates. *AIDS and Bahavior, 12,* 321.

Holtzman, D., Chen, S., Zhang, S. et al. (2003). Current HIV/AIDS-related knowledge, attitudes and practices among the general population in China: Implications for action. *AIDScience, 3*(1), 6.

Hong, Y., Li, X., Fang, X. et al. (2009). HIV/AIDS-related risks and migratory status among female sex workers in a rural Chinese county. *AIDS Care, 21*(2), 212.

Jeffries, E. (2006). *Sex and sexuality in China*. Oxon, UK: Routledge.

Jie, S., Kangmai, L., Menjie, H. et al. (2004). The Chinese HIV/AIDS epidemic and current response. In *AIDS in Asia* (47). New York: Kluwer Academic Press.

Kalichman, S., & Rompa, D. (1995). Sexual sensation seeking and sexual compulsivity scales: reliability, validity and predicting HIV risk behaviour. *Journal of Personality Assessment, 65*(3), 586.

Kallings, L. (2008). The first post-modern pandemic: 25 years of HIV/AIDS. *Journal of Internal Medicine, 263*(3), 218-243.

Kaufman, J., & Jing, J. (2002). China and AIDS - the time to act is now. *Science, 296*(5577), 2339-2340.

Lagarde, J., Schim van der Loeff, M., Enel, C. et al. (2003). Mobility and the Spread of Human Immunodeficiency Virus into Rural Areas of West Africa. *International Journal of Epidemiology, 32,* 744-752.

Laquian, A. (1996). Rural-urban transition and the planning and governance of mega-urban regions. In *The rural-urban transition and development in China* (Proceedings of an international conference, Zhongshan University, Guangzhou, China, 93-114). Vancouver: Center for human settlements, University of British Columbia.

Li, Q. (1999). Effects of labour out-migration on income growth and inequality in rural China. *Development and Society, 28*(1), 93.

Li, Q. (2001). Zhongguo waichu nongmingong ji qi huikuan zhi yanjiu (Research on China's rural out-migrants and their remittances). *Shehuixue yanjiu (Sociological Research), 4,* 64.

Lieber, E., Li, L., Wu, Z. et al. (2006). HIV/STD stigmatization fears as health-seeking barriers in China. *AIDS and Behavior, 10*(5), 463.

Lönn, E., Sahlholm, K., Maimaiti, R. et al. (2007). A traditional society in change encounters HIV/AIDS: knowledge, attitudes and risk behavior among students in Northwestern China. *AIDS Patient Care and STIs, 21*(1), 48.

Lu, Y., Essex, M. *AIDS in Asia*. New York: Kluwer Academic Press, 2004.

Mattson, C., Campbell, R., Karabatsos, G. et al. (2008). Scaling sexual behaviour or "sexual risk propensity" among men at risk for HIV in Kisumu, Kenya. *AIDS and Behavior, online.*

McClelland, G., Teplin, L., Abram, K. et al. (2002). HIV and AIDS risk behaviours among female jail detainees: implications for public health policy. *American Journal of Public Health, 92*(5), 818.

Murphy, R. (2006). Domestic Migrant Remittances in China: distribution, channels and livelihoods. *IOM migration research series* (24), Geneva: Switzerland: International Organisation for Migration.

Osgood, D., McMorris, B., & Potenza, M. (2002). Analysing measures of crime and deviance I: item response theory scaling. *Journal of Quantitative Criminology, 18*(3), 1573.

Qian, Z., Vermund, S., & Wang, N. (2004). Risk of HIV/AIDS in China: subpopulations of special importance. *Sexually Transmitted Infections, 81,* 442-447.

Rao, D., Pryor, J., Gaddist, B. et al. (2006). Stigma, secrecy and disrimination: ethnic/racial differences in the concerns of people living with HIV/AIDS. *AIDS and Behavior, 12,* 265.

Rongyao, H., Yuan, Y., & Li, T. (2000). Study on reproductive health among unmarried women in floating population. *Chinese Journal of Family Planning, 8*(5), 212-214.

Rowland, D. (1992). Family characteristics of internal migration in China. *Asia Pacific Population Journal, 7*(1), 3-26.

Schafer, J., & Graham, J. (2002). Missing data: our view of the state of the articulating. *Psychiological Methods, 7*(2), 147.

Schultz, W., & Sibberns, H. (2004). IEA Civic Education Study Technical Report. In Schultz, W.; Sibberns, H., *IEA civic education study technical report* Amsterdam, Netherlands: IEA Secretariat.

Shao, Y., & Lv, F. (2004). Current HIV/AIDS diagnosis, reporting and surveillance in China. In *AIDS in Asia* (351). New York: Kluwer Academic Press.

Shao, Y. (2006). AIDS epidemic at age 25 and control efforts in China. *Retrovirology 3,* 87

Sheeran, P., & Taylor, S. (1999). Predicting intentions to use condoms: a meta-analysis and comparison of the theories of reasoned action and planned behavior. *Journal of Applied Social Psychology, 29*(8), 1624-1675.

Shisana, O., & Simbayi, L. (2002). *Nelson Mandela HSRC Study of HIV AIDS: HSRC Study of HIV/AIDS*

Full Report. Cape Town: HSRC Press.

Smith, C. (2005). Social geography of sexually transmitted diseases in China: exploring the role of migration and urbanisation. *Asia Pacific Viewpoint, 46*(1), 65-80.

S-Plus. (Software). (2006). *S-Plus Workbench User's Guide*. Seattle, USA: Insightful Corporation.

SPSS Inc. (Software). (2007). *Statistical Package for the Social Sciences*. Chicago: SPSS Inc.

Stack, S. (1994). The effect of geographic mobility on pre-marital sex. *Journal of Marriage and the family, 56,* 204-208.

Stares, S. (2008). *Latent trait and latent class models in survey analysis: case studies in public perceptions of biotechnology.* PhD, London School of Economics.

UNAIDS. (2005). Monitoring the Declaration of Commitment on HIV/AIDS: guidelines on construction of core indicators. (Geneva, Switzerland), UNAIDS.

UNAIDS. (2007). Guidelines on construction of core indicators. In De Lay, P., *Monitoring the declaration of commitment on HIV/AIDS* (7).

UNAIDS/WHO. (2008). Asia: AIDS epidemic update: regional summary. *Asia: AIDS epidemic update: regional summary* Geneva, Switzerland: UNAIDS/ WHO.

Vermunt, J., & Magidson, J. (Software package). (2005). *Latent Gold 4.0* (Statistical Innovations, Inc.). Belmont, MA:.

Wang, B., Li, X., Stanton, B. et al. (2007). Gender differences in HIV-related perceptions, sexual risk behaviors, and history of sexually transmitted diseases among Chinese migrants visiting public sexually transmitted disease clinics. *AIDS Patient Care and STDs, 21*(1), 57.

Wang, B., Li, X., Stanton, B. et al. (2008). Correlates of HIV/STD testing and willingness to test among rural-urban migrants in China. *AIDS and Behavior, Online First,* 10.1007/s10461-008-9482-1.

Wang, D., & Cai, F. (2007). Migration and poverty alleviation in China. *IOM Migration research series* (27), Geneva: Switzerland: International Organisation for Migration.

White, R. (2003). Commentary: what can we make of an association between human immunodeficiency virus prevalence and population mobility?. *International Journal of Epidemiology, 32,* 753.

WHO. (2005). Sexual and reproductive health of adolescents and youths in China - a review of literature and projects 1995-2002. *WHO regional office for the Western Pacific* Manila, Philippines: WHO.

Wilkinson, R., & Pickett, K. (2005). Income inequality and population health. *Social science and medicine, 62,* 1768-84.

Wilson, W. (1987). *The truly disadvantaged: the inner city, the underclass, and public policy.* Chicago: University of Chicago Press.

Xiao, Y., Kristensen, S., Sun, J. et al. (2006). Expansion of HIV/AIDS in China: lessons from Yunnan Province. *Social Science and Medicine, 64,* 665.

Yang, H., Li, X., Standon, B. et al. (2005). HIV-related risk factors associated with commercial sex among female migrants in China. *Health Care Women International, 26,* 134-148.

Yang, L., & Kleinman, A. (2008). 'Face' and the embodiment of stigma in China: the cases of schizophrenia and AIDS. *Social Science and Medicine, 67,* 398-408.

Yang, X. (2006). Temporary migration and HIV risk behaviors in China. *Environment and Planning A, 38,* 1527-43.

Yang, X., Derlega, V., & Luo, H. (2007). Migration, behaviour change and HIV/STD risks in China. *AIDS Care, 19*(2), 282-288.

Yang, X., & Xia, G. (2006). Gender, migration, risky sex and HIV infection in China. *Studies in Family Planning, 37*(4), 241-250.

Yang, X., & Xia, G. (2008). Temporary migration and STD/HIV risky sexual behavior: A population-based analysis of gender differencesin China. *Social Problems, 55*(3), 322-346.

Yi, Z., Hua, X., & Jiaxi, Z. (2004). Infectious diseases in China. In *AIDS in Asia* (295). New York: Kluwer Academic Press.

Yinhe, L. (2006). Regulating male same-sex relationships in the People's Republic of China. In *Sex and sexuality in China* (82-101). Abingdon, UK: Routledge.

Yu-Hua, W., Li-Shi, B., Yan-Cheng, L. et al. (2004). Men who have sex with men and HIV/AIDS prevention in China. In *AIDS in Asia* (365). New York: Kluwer Academic Press.

Zhongwei, Z. (2006). Income inequality, unequal health care access, and mortality in China. *Population and Development Review, 32*(3), 461-483.

Zhou, J., Sun, X., Mantell. J. et al. (2007). AIDS-related knowledge, attitudes and behavior survey among the "migrant" population in China. *Journal of reproduction and contraception, 18*(2), 155.