School of Business
Unlimited Future, Unlimited Possibilities
Te Kura Pakihi

ISSN 0111-1760 (Print)
ISSN 1178-2293 (Online)

University of Otago
Economics Discussion Papers
No. 0806

July 2008

The determinants of international students' return intention

Jan-Jan Soon

Correspondence to:
Jan-Jan Soon
Department of Economics
University of Otago
PO Box 56, Dunedin
New Zealand 9054
Phone: (64) 3 479 4085
Fax: (64) 3 479 8174
Email: j-jsoon@business.otago.ac.nz

Abstract

Students' non-return is a specific type of brain drain. This paper is an empirical study of the determinants of students' return intention in New Zealand. Applying a binary logit model on a comprehensive set of survey data, this study finds that initial intention prior to leaving for abroad is the most important factor determining whether or not a student intends to return home after completing his tertiary education. Students' perceptions on comparative aspects of the home and host country, such as wage competitiveness, working environment, opportunities for knowledge application and lifestyle, also contribute significantly to return intention.

JEL Classification: C25, J61

Keywords: Students' non-return, brain drain, binary logit model

Overview

Student brain drain cannot be ignored in the analysis of brain drain (Vinokur, 2006, p. 19). This paper contributes empirically to the study of the determinants of students' return intention in New Zealand. It concludes that initial intention prior to coming to New Zealand is the most important factor determining whether or not a student intends to return home after completing his tertiary education. Perception-related variables on comparative aspects of the home and host country, such as wage competitiveness, working environment, opportunities for knowledge application, lifestyle and the network of family and friends, also contribute significantly to return intention.

The non-return of students to their home countries following study abroad is one type of brain drain; brain drain itself is a form of migration. According to Becker (1962; 1964) and Sjaastad (1962), migration is a manifestation of human capital investment. In Becker's human capital theory, an individual would invest in human capital if the benefits of his doing so outweigh the costs, be it by studying for a degree, or upgrading skills. Similarly, in Sjaastad's human capital theory of migration, an individual would contemplate migrating if the real or perceived benefits from the move were higher than the costs incurred.

The brain drain phenomenon has been observed since the 1950s when there was an exodus of scientists from Britain to the United States. A long-running debate over the brain drain issue emerged in the 1960s. The main protagonists in that era were Grubel and Scott (1966) who supported the nationalist[1] view of perfectly competitive markets,[2] while Johnson (1967) defended the internationalist[3] argument pertaining to the effects of brain drain. In the 1970s, Bhagwati and his colleagues (J. Bhagwati & Hamada, 1974; J. Bhagwati & Rodriguez, 1975; J. N. Bhagwati & Dellalfar, 1973; Hamada & Bhagwati, 1975) highlighted the harmful effects of brain

[1] A concept that visualizes economic and cultural welfare in terms of the welfare of the residents of a nation, viewed as a totality (Johnson 1967, p. 379).
[2] In their view, wages always equal marginal product; thus, there would be no negative welfare impact whatsoever on those left behind (TLB). TLB is a term coined to refer to the natives of a country who did not emigrate.
[3] The view that the international circulation of human capital is a beneficial process (Johnson 1967, p.380).

drain during that period and even proposed taxing the brain drain.[4] Such a tax has never been implemented (Lowell & Findlay, 2001).

These early models of detrimental brain drain effects are complemented by more recent models, with the beneficial brain drain (BBD) model[5] being the most notable. Its human capital formation hypothesis has received much recent attention (Beine, Docquier, & Rapoport, 2001, 2003; Mountford, 1997; Stark, Helmenstein, & Prskawetz, 1997, 1998; Stark & Yong, 2002; Vidal, 1998). Schiff (2005), however, has claimed that the brain gain has been overestimated.

Theoretical studies on the determinants of brain drain were pioneered by Kwok and Leland's (1982) seminal paper, in which they show that asymmetric information on the part of the home country employers is one of the factors contributing to Taiwanese students' non-return in the United States. Subsequent works in a theoretical vein have identified a number of other factors pertaining to brain drain, such as signalling (D.-H. D. Lien, 1987),[6] scale economies in education (Miyagiwa, 1991), on-the-job training (Chen & Su, 1995), learning-by-doing (Wong, 1995), inappropriate type of research (D.-H. D. Lien, 1988) and national (D. Lien, 2008) and international (D. Lien, 2006) accreditation requirements.

In the theoretical pursuit of elegant theories to explain brain drain and students' non-return, the empirical application of such theories has taken a backseat. Students' non-return is a subtle yet prevalent form of brain drain. It has been a perennial concern for most developing countries. Studies on students' non-return have concentrated on the United States; the few notable empirical studies include those by Gungor (2003), Bratsberg (1995) and Huang (1988).

There have been a few attempts to look at brain drain in New Zealand (Brown & Connell, 2004; Gani & Ward, 1995), but these have focused primarily on high-skilled immigration and not students' non-return specifically. Students' non-return is a

[4] Also known as the Bhagwati tax.
[5] Beneficial in terms of creating incentives for human capital formation resulting in higher average education levels in the home countries, in anticipation of migration possibilities.
[6] Lien extends Kwok & Leland's paper by adding a signalling component into the asymmetric information framework.

non-trivial component of brain drain that deserves a closer look. It is this gap in the brain drain literature which this paper attempts to fill.

The purpose of this paper is to find out and examine closely the principal explanatory factors in predicting students' return probabilities. Their importance will also be evaluated using a range of interpretations, such as marginal effects and odds ratios.

The remainder of the paper is structured as follows. The next section discusses the data, followed by a brief review of the theoretical method and a section on empirical model specification. After describing the variables employed for this paper, the main results and discussions follow. The final section concludes.

Data

Although the New Zealand Department of Labour has produced a report (Merwood, 2007) on the retention of international students, it nevertheless falls short for the purpose of empirical work, being a descriptive study of all international students in New Zealand.[7] Due to the lack of data for the paper's specific objectives, this paper generates its own microdata through a web-based questionnaire survey. As pointed out by Winkelmann and Boes (2006, p. 3), the defining feature of microdata lies in their cross-sectional, observational and discrete nature. Microdata pertain to individual-level decision-making issues (Heckman, 2000).

The target population of this study is the group of international students currently studying in all eight universities in New Zealand. A questionnaire was prepared to be sent out to them via the universities international offices. However, only two[8] universities agreed to let their international students participate in the web-based survey.[9] The sampling frames used for this paper are comprehensive[10] and there is no need to construct self-compiled sampling frames,[11] which are often not exhaustive. Short-term exchange students are excluded from the survey.

The questionnaire solicits information on demographic variables, individual-specific attribute variables and perception-related variables on either a home, host or a third country. The response rates[12] are approximately 29.6% and 14.7% from the University of Otago and University of Canterbury respectively. The data are then inspected for missing responses. The final usable sample size comes to 516 respondents.[13] Only the non-bonded (by scholarship) sub-sample is used.

---

[7] International students who are in New Zealand attending high schools, private training establishments, English language schools, polytechnics, institutes of technology, teachers' training colleges and universities.
[8] University of Otago and University of Canterbury.
[9] For the purpose of this paper, which is part of a broader doctoral research, the sample used comes from these two universities only.
[10] The sampling frames used for this paper are the lists of all international students as maintained by the universities' International Offices.
[11] Such as those by Gungor (2003) and Kao & Lee (1973).
[12] These are not the final response rates as the survey is still ongoing at the time this paper is written.
[13] As Long (1997, p. 54) has recommended, a sample size over 500 seems adequate for maximum likelihood estimation.

Table 1 shows the respondents' demographic profile. The respondents are divided into four age groups – under 20 years old, between 20 and 24, between 25 and 29 and over 29. The majority of the respondents are from the 20 to 24 age group, constituting almost 50% of the total respondents. The number of single and married respondents is disproportionate, with over 90% of them being single. However, there is a balance between single males and single females.

Precautions have been taken to ensure the validity and rigour of the research. In terms of data collection, saturation sampling is used, where the web-based questionnaire is disseminated to every international student (excluding short-term exchange students) on the lists of the universities' international offices. In this sense, saturation sampling is unlike conventional sampling designs, and is more like a census (Sue & Ritter, 2007). After the data collection phase, measures have also been taken to check the data for inconsistencies and discrepancies. A handful of data entry errors were identified and removed.

Brief review of theoretical method

The objectives of this paper, and the type of disaggregated micro-level data used, dictate its empirical method: the binary logit model. This model comes from the family of probability models, better known as discrete choice models (DCM). At the heart of probability models is random utility theory, or RUT (Manski, 1977).[14] According to RUT, an individual is capable of evaluating the utility associated with a set of viable alternatives and subsequently selecting the alternative that he perceives will yield maximum utility.

RUT incorporates both deterministic and stochastic elements, reflecting the observable and unobservable attributes of the alternatives. Note that the unobservable attributes of the alternatives are only relevant to the researcher, and not to the decision-maker. According to RUT, a decision-maker *i* attaches a certain utility level to each alternative *k* in his choice set, i.e., $U_{ik} = v_{ik} + \varepsilon_{ik}$, with v_{ik} being the deterministic and ε_{ik} the stochastic component respectively. He would choose the

[14] Manski (1977) gives a formal exposition of the random utility theory.

alternative with the highest utility, in which case the variable of interest would be $P(U_{ik} > U_{ij}) \quad \forall j \neq k$. In order to compute probabilities such as $P(U_{ik} > U_{ij})$, assumptions have to be made for the distribution of the error terms (i.e., ε_{ik}, $\varepsilon_{ij} \forall j \neq k$). Here, it is assumed that the error terms[15] are identically and independently distributed according to the Gumbel distribution.[16] The difference between the error terms ($\varepsilon_{ik} - \varepsilon_{ij}$) then has a logistic distribution, which leads to the logit model.[17]

DCM first originated in the field of transportation economics (Ben-Akiva & Lerman, 1989; Domencich & McFadden, 1975; K. Train, 1986). Recently, DCM has become fairly standard in brain drain, students' non-return and migration analyses. DCM operates at the individual decision-makers' level (K. E. Train, 2003, p.33). DCM is used to model selection from within a discrete set of exhaustive alternatives. The binary logit model, a member of the DCM family, has two competing alternatives.

Empirical model specification

This paper uses the binary logit model estimated using the maximum likelihood technique.[18] This model addresses the dichotomy of whether or not a student intends to return home and identifies the determinants of such intentions. The model is set out formally as follows.

$$\text{Let} \quad y_i = \begin{cases} 1 \text{ if a student } i \text{ intends to return} \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

[15] There is one disturbance per individual per choice.
[16] Otherwise known as Type 1 Extreme Value distribution. Also note that a Gumbel distribution is obtained from the log of a Weibull distribution.
[17] Its cumulative distribution function, cdf, of the error terms by the Gumbel or Type 1 Extreme Value distribution is $F(\varepsilon_{ij}) = \exp(-\exp^{-\varepsilon_{ij}})$ and its equivalent probability distribution function, pdf, is $f(\varepsilon_{ij}) = \exp^{-\varepsilon_{ij}} \exp(-\exp^{-\varepsilon_{ij}})$.
[18] The maximum likelihood estimator is asymptotically consistent, normal and efficient.

By binary logit, the predicted probability of return is then,

$$\hat{P}(y_i = 1 | X) = \frac{e^Z}{1 + e^Z} = F(Z) \quad (2)$$

where

$F(Z)$ is the cumulative distribution function

and $$Z = \hat{\beta}_0 + \hat{\beta}_1 \mathbf{X}_1 + \hat{\beta}_2 \mathbf{X}_2 + \hat{\beta}_3 \mathbf{X}_3 \quad (3)$$

in which

$\mathbf{X}_1$ = vector of demographic variables

$\mathbf{X}_2$ = vector of individual-specific attributes

$\mathbf{X}_3$ = vector of perception variables

$\hat{\beta}_k$ = estimated parameter coefficients

Although there is no competing model proposed in this paper, the model presented is nonetheless tested for its specification fit using the ‘linktest’ method. As shown in Table 2, the hat squared is found to be insignificant (i.e. a p-value of 0.592), indicating that there is no model mis-specification. This is because, if the model suffers no mis-specification, the hat squared variable should not have any predictive power except by chance. Essentially, the ‘linktest’ method uses the predicted value (i.e. ‘_hat’ in Table 2) and its squared (i.e. ‘_hatsq) as the predictors in reconstructing the model. Simply put, this method is used to detect if the model has included all relevant regressors and excluded irrelevant ones.

Variable choice and description

Table 3 contains the list of variables with their descriptions. The dependent variable is whether or not a student currently intends to return home after completing their studies here in New Zealand. Equation (1) from the previous section summarizes the binary nature of the dependent variable. In constructing this dependent variable, a respondent is asked their current return intention, ranging from definitely return, probably return, probably not return and definitely not return. In order to estimate a binary logit model, the responses are then collapsed into two categories – return or not return.

The choice of explanatory variables is largely based on fundamental economic theories,[19] brain drain and students' non-return theories, hypotheses and findings from the literature. Attempts have been made to exclude arbitrary factors from the so-called 'push-pull model' (i.e. retention versus repulsion factors) developed by Lee (1966) as they are often not exhaustive and subject to the researcher's preference.

As mentioned in the previous section, this paper employs three different sets of explanatory variables to explain the students' return intentions. The set of demographic variables consists of gender, marital status, age and a squared term of age. The squared term is intended to capture diminishing or increasing effects of an additional year of age on the predicted return probabilities. This set of demographic variables is almost always used in the empirical literature.

The second set of explanatory variables includes other individual-specific attributes, such as socio-economic background, level and year of study, years of residence in New Zealand and years of prior professional working experience. Squared terms of both the years of residence in New Zealand and years of working experience are included to allow for any reverse trends in the predicted return probabilities. A measure of mobility is also used, as proxied by whether an individual has studied somewhere else[20] before enrolling for their current study. The fields of

[19] The interplay among consumer behaviour theory, utility maximization theory and the human capital theory of migration.
[20] Other than their home country.

study are included in order to test Chen and Su's (1995) hypothesis that students' non-return is higher in capital-dependent[21] disciplines.

The final set of explanatory variables is composed of perception-related variables, measuring how respondents perceive different aspects of either their home country or the country in which they intend to reside; for example, perceptions of wages, working environment, opportunities to apply knowledge, lifestyle, family ties, network of friends and race equality.

The variable on the perception of wages measures how a student perceives the wage competitiveness in his home country as compared to foreign wages. It is postulated that wage competitiveness is one of the main factors in determining a student's return intention. The variable on the perception of working environment is in terms of good physical working conditions, competent colleagues and adequate research facilities. As for the perception of the opportunity for knowledge application, it measures how a student perceives he can utilize his knowledge and practically apply it. Certain fields of knowledge may find less application in a less developed home country.

The lifestyle perception variable looks at how a student perceives the lifestyle in his home country or the lifestyle in a foreign country. If say, a student perceives the lifestyle to be better abroad, then this perception may play a role in shaping his non-return intention. The variable on the perception of family ties is to measure how strong a student perceives the family bond and network of friends. A perception of closer family ties at home would influence one's return intention. As for the race equality perception variable, it looks at how a student perceives race equality at home as compared to a foreign country.

Results and discussion: On coefficient signs and significance

From the logit estimation in Table 4, all the explanatory variables have the expected signs and 11 of them are significant. In logit models, the magnitude of the

[21] A discipline or profession that needs a large stock of physical and human capital, without which it cannot function.

estimated coefficients does not have the usual interpretation as its OLS counterpart. Instead, more meaningful interpretations would be in terms of predicted probabilities, marginal effects and odds ratios which will be discussed below. Note also that the model is estimated using a subsample where only those who are not bonded (by scholarships) are included. This is to avoid any bias that may arise, as bonded students are those who simply have to return regardless of their own wishes.

The set of demographic variables – gender, age and marital status – is found to be insignificant. This may be due largely to the homogeneity of the respondents, especially in terms of age and marital status: approximately 50% of the respondents are in the 20 to 24 age group, and over 90% of the respondents are single. This lack of variation may be part of the reason why the age and marital variables are insignificant. However, such insignificance should not be a cause for concern, as the empirical evidence in the previous literature is mixed and inconclusive as well. Furthermore, there are no theoretical underpinnings dictating the signs of these demographic variables.

With the exception of the perception variable on race equality, all the other five perception-related variables are significant at the 5% level. They all have the expected positive signs, indicating that the return probability increases when the home country is perceived to be better in terms of competitive wages (goodHwage), working environment (goodHwenv), opportunities to apply specialized knowledge (goodHoppk), preferred lifestyle (goodHlife) and strong network of family and friends (goodHfren).

The residence years in New Zealand variable (i.e. ‘yrsinNZ’) is significant at the 10% level and has the expected negative sign. The longer a student resides in the host country, the more his return probability decreases. The longer one stays in a host country, the more one assimilates to its culture and society (Gungor, 2003; Kao & Lee, 1973; Waldorf, 1995). Furthermore, one may face adjustment difficulties when one returns home. There is, however, no evidence that the probability trend will take an opposite turn, as the squared term of the residence years (i.e. the ‘yrsinNZsq’ variable) is insignificant.

Being a doctoral student is found to be significant at the 10% level. The expected negative sign implies that being a doctoral student decreases the return probability. Having a higher level of human capital gives the student more potential destination options to choose from. As pointed out by Lien (1988), a doctoral student usually engages in a high-income type of research which is associated with more developed countries. Upon graduation, such a student may be reluctant to return home as there would be fewer opportunities for him to continue with his high-income type of research.[22] This is supported by the data: only about 22% of doctoral students (i.e. 29 out of the 131 doctoral students in the sample) perceive good opportunities to apply their knowledge at home.

The health science discipline is found to be significant at the conventional 5% level, compared to the humanities discipline. The negative sign of the coefficient indicates that students in the health science discipline are less likely to return home compared to those from humanities. This may be due to the possibility that the health science discipline utilizes more capital than the humanities discipline. As theorized and concluded empirically by Chen and Su (1995), those studying in more capital-dependent disciplines are less likely to return. However, in this paper, this cannot be concluded for the science discipline as it is insignificant in relation to the humanities discipline. This may be due to the fact that the science discipline itself can be further divided into many sub-disciplines, each with quite different labour market conditions.

The initial return intention variable is highly significant at less than 1% significance level. The positive coefficient means that a student having an intention of returning home prior to his coming to New Zealand would have a higher return probability. This positive relationship has also been observed in the previous students' non-return literature (Gungor, 2003; Myers, 1973; Zweig, 1997). Table 5 shows the sample proportions between this variable and the dependent variable, where approximately 83% of the respondents whose initial intention was to return, have indicated a current return intention.

[22] This issue has been identified by Grubel and Scott (1977, p. 148) who suggest that foreign students be discouraged from working on highly sophisticated research which tends to over-educate them with respect to the needs of their home countries.

The remaining two significant variables pertain to a student's family support and background. Family support of one's non-return decision is also found to be highly significant. The negative sign shows that if one's family is supportive of a student's non-return decision, then his return probability decreases. The father's education variable, capturing a student's socio-economic background, is significant at the 10% level. The result here shows that students from a higher socio-economic status are more likely to return home. Consequently, this may imply that students from a lower socio-economic background are the ones who would be most likely to stay on or move to another third country.

Due to the complexities of nonlinear models such as a binary logit model, there are other possible interpretations than just the expected signs and significance. Unlike its OLS counterpart, where most of the interpretation work is completed once the expected signs and significances are obtained, in logit models many subsequent interpretations can be derived (Long, 1997, p. 61). Recall from equation (2) that one of the main purposes of the model is to predict return probabilities. The remaining sections discuss other interesting interpretations of the model, including predicted probabilities, marginal effects and odds ratios.

Results and discussion: On predicted probabilities

From the analysis, an 'average' student has a predicted probability of 0.5997 of returning home. 'Average' here is in terms of all the explanatory variables taking their mean values. However, this figure does not make much sense as most of the explanatory variables are in dummy form, where mean values do not have any real meaning. Thus, the predicted probabilities would be of more interest if estimates were based on certain hypothesized scenarios. Also note that in Figure 1, which shows the cumulative distribution function of the predicted probability, the distribution takes on a non-linear, sigmoid-shaped relationship between the independent variables and the predicted probability.[23]

[23] When the outcome variable is a probability, the independent variables would have diminishing effects as the predicted probability → 0 or → 1 (Long 1997, p. 39).

In Table 6, three different hypothesized scenarios are used to illustrate the predicted probability of returning. To begin with, a baseline scenario (i.e. Scenario One) is constructed so that all the continuous variables are set at their mean values, while insignificant categorical variables are set at their modal values. Scenario One depicts a student from a relatively developed country as indicated by the perception-related variables.[24] Such a student has a good socio-economic family background and his prior intention is to return home. All his other attributes are as shown in Table 6. In this scenario, the predicted return probability is a high 0.9871.

The remaining four hypothesized scenarios are changed in relation to the baseline Scenario One. For example, Scenario Two also depicts a student from a developed country, but who comes from a family of lower socio-economic background. This final-year health science doctoral student has an initial non-return intention. Combined with other attributes, his predicted return probability is now 0.7611.

Scenario Three is similarly interpreted. In this scenario, a student comes from a developing country.[25] He has no prior intention of returning and he comes from a less well-to-do family. This gives him a low 0.098 predicted return probability. This simply means that his predicted probability of returning home is less than 10%. This, however, is a somewhat extreme case.

Scenarios Four and Five are basically the same as Scenario Three, except that the residence years and working years are altered to see how the predicted return probabilities change. Scenario Four depicts a student who has been in New Zealand for more than a decade (the maximum residence years in the sample is approximately 11.25) but has no previous working experience in his home country. His predicted return probability is 0.1827.

---

[24] This paper presumes that the perception-related variables would take on a value of 1 (which means that a certain aspect is better at the home country) for someone from a more developed country. The dummy variable to represent developed and developing countries is not included in the model as it gives a wrong expected sign.

[25] Where the perception-related variables, with the exception of the perception on the network of family and friends, would take on a value of 0. Such zero values indicate less favourable conditions which are salient features of a typical developing country.

On the other hand, Scenario Five is a reverse of Scenario Four. A student now has years of professional working experience (the maximum working years in the sample is 18) but has resided for just slightly less than a year in New Zealand. His predicted return probability increases to 0.6605.

Results and discussion: On marginal effects

Perhaps of even more interest are the partial changes in the predicted probabilities (i.e. the marginal change of a continuous variable or the discrete change of a categorical variable on the predicted return probabilities, collectively referred as marginal effects in this paper).[26] One caveat to note here is that, in nonlinear models such as binary logits, the effect of a change in a particular explanatory variable on the predicted probability would depend on the specific levels of all other explanatory variables in the model. In addition to that, the marginal effect of a particular variable of interest also depends on its initial level as well as the magnitude of its change. Thus, the marginal effects are non-constant as compared to those of the OLS models.

Before delving further into the marginal effects results, it is a good idea to have a brief look at the change in predicted probabilities as shown in Table 7. It shows the extent to which changes in a certain explanatory variable affect the predicted probability, by letting the variable vary from its minimum to maximum value, while holding other variables at their fixed values. For a categorical variable, its minimum value is zero while the maximum is one. All continuous variables are set at their mean values, while all categorical variables are at their modal values.[27]

The last column in Table 7 gives the range of the changes in predicted probabilities. As noted by Long (1997, p. 66), more can be learnt by analyzing variables whose ranges of predicted probabilities are relatively large. The residence years, initial intention and lifestyle perception variables have large ranges of predicted

[26] Also known as marginal probability effects (Winkelmann & Boes 2006).
[27] As noted before, it is less meaningful to set categorical variables at their mean values. This is an alternative suggested by Long (1997, p. 77).

probabilities. For example, the predicted probabilities for the residence years variable range from 0.6419 when a student has resided slightly less than a year in New Zealand to a probability of 0.0543 when the residence years are about 11. The corresponding negative difference of 0.5875 indicates the decrease in probabilities.

In Table 8, the marginal effects are given in the dy/dx column, where the signs correspond with those of the coefficients in Table 4. Focusing on significant variables, the initial return intention categorical variable has the largest marginal effect on the predicted return probabilities. Holding all other variables at their fixed (i.e. either at modal or mean) values, when a student initially intends to return, his return probability increases by approximately 0.43. Otherwise, his predicted return probability drops by 0.43, ceteris paribus.

The marginal effects of all the significant perception-related variables can be interpreted in a similar fashion. For example, if a student perceives that the lifestyle is better at home, his return probability increases by 0.37, all other variables held at their fixed values. Indeed, lifestyle in the home country is a strong determinant, which has been recognized since Kwok and Leland’s (1982) seminal paper, where they include a discount factor for those who choose to work abroad, in representing the preferred home country lifestyle.

Note that the marginal effect of the perception on competitive wage (i.e. ‘goodHwage’) is the smallest of the significant effects. This result is compatible with the findings in the brain drain (Agarwal & Winkler, 1984; Brown & Connell, 2004; Gani & Ward, 1995) and students’ non-return (Kao & Lee 1973; Huang 1988; Bratsberg 1995; Gungor 2003) literature where wage does not seem to play a dominant role. For high-skilled individuals, there are many factors other than wages that determine their return intentions. In this instance, they perceive lifestyle, a network of family and friends, the working environment and opportunities for knowledge application – in that order – to be more important than the wage.

As for the residence years variable (i.e. ‘yrsinNZ’), which is continuous, each additional year of residence in New Zealand decreases the return probability by about 0.076. This change, although small, is non-trivial. The longer one resides in a host country, the less likely one would return home. Figure 2 shows the marginal effect of residence years on the predicted return probabilities. The probability of returning decreases the longer one resides in the host country (i.e., New Zealand).

Results and discussion: On odds ratio

An advantage of the logit model over the probit model is that it lends itself to odds ratio interpretations. There are two ways of interpreting odds ratios – in terms of factor change[28] and percentage change. They convey the same information, basically the effect of a particular variable on the odds of a positive outcome occurring. Table 9 shows the odds ratio of returning (positive outcome) versus not returning (negative outcome).[29]

In Table 9, there are two outstanding odds ratios. The odds of returning home are 8.4 times larger when one expresses an initial intention of returning home, while holding other variables constant.[30] Similarly, having an initial intention to return increases the odds of returning by about 740%. Having a favourable perception on lifestyle in one's home country increases the odds of returning by 468%.

All other significant perception-related variables have relatively high odds ratios. For each additional residence year in New Zealand, the odds of returning home decrease by 26.5%, or 0.7 times smaller. Remaining significant variables are similarly interpreted. Their effects, though modest, are non-trivial. Take, for instance, a doctoral student – the odds of returning home decrease by about 45%, compared to non-doctoral students.

Summary and conclusion

[28] A factor change (or odds ratio) of more than unity corresponds with the positive coefficient sign and vice versa.
[29] It is typical to consider the odds of observing a positive outcome compared with a negative one (Long & Freese 2003, p. 145).
[30] The interpretation of odds ratio assumes that the other variables are held constant, but it does not require that they be held at any specific values (Long & Freese 2003, p. 147).

This paper sets out to investigate the factors governing the return intention of international students who are currently studying in New Zealand universities. A number of principal[31] determinants have been identified. They are the initial return intention, and most of the perception-related variables. The perception on lifestyle, however, stands out among the perception-related variables.

Having a return intention prior to coming to a host country has the greatest positive impact on a student's current return intention, as shown in its marginal effect (approximately 0.43) and odds ratio factor (approximately 8.4). However, this is not to say that the prior intention never changes. Approximately 36% of the respondents' current intentions are different from their initial intention.

Five out of the six perception-related variables are significant (at either 1% or 5% level). The perception of wage competitiveness does not have a large impact when compared to the perceptions on preferred lifestyle, network of family and friends, working environment and opportunities for application of one's specialized knowledge. The perception of race equality, however, does not seem to contribute to one's return intention (i.e., the 'goodHrace' variable is insignificant).

The other significant variables in explaining students' non-return are residence years in New Zealand (or the education host country), studying at a doctoral level, discipline of study (the health science discipline in particular), family support of one's non-return intention and family socio-economic background (as proxied by the father's education level).

The combination of all the significant variables contributes to the current return intention of the students. The identification of principal variables is important for policy-making purposes, either by the home or the host country. Inevitably, the interests of the home and host country would be in competing directions, with the home country aiming to attract the return of their students, while the host country

[31] In terms of relatively large marginal effect and odds ratio.

trying to retain the bests. Thus, understanding the underlying factors of the students' intentions would help make better policies.

Table 1: Respondents' demographic profile

Age group	Single		Married		Total
	Female	Male	Female	Male	(%)
Below 20	28	26	0	0	54 (10.5)
20 to 24	129	122	1	3	255 (49.4)
25 to 29	51	68	9	12	140 (27.1)
Above 29	24	18	10	15	67 (13.0)
Total	232 (45.0)	234 (45.3)	20 (3.9)	30 (5.8)	516 (100.0)

Table 2: Model mis-specification test

Logistic regression			Number of obs = 516			
			LR chi2(2) = 239.40			
			Prob > chi2 = 0.0000			
Log likelihood = -236.08521			Pseudo R2 = 0.3364			
binaryreturn	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
_hat	1.01513	.0928704	10.93	0.000	.8331071	1.197152
_hatsq	-.0251967	.0469922	-0.54	0.592	-.1172998	.0669064
_cons	.0427932	.1395554	0.31	0.759	-.2307303	.3163166

Table 3: Description of variables

Variable	Variable descriptions
Dependent variable	
binaryreturn	return intention 1 = intend to return, 0 = otherwise
Demographic variables	
q2_age	age
agesq	squared term of age
male	gender, 1 = male, 0 = otherwise
single	marital status, 1 = not married, 0 = otherwise
Other individual-specific attribute variables	
yrsinNZ	duration of residence years in New Zealand
yrsinNZsq	squared term of yrsinNZ
workyrs	professional working experience in terms of years
workyrssq	squared term of workyrs
hselsewhr	have had high school education in a country other than home 1 = have studied elsewhere, 0 = otherwise; a proxy of mobility
phd	currently studying for a doctoral degree, 1 = yes, 0 = otherwise
finalyear	1 = final year students, 0 = otherwise
commerce	1 = in the commerce discipline, 0 = otherwise
hscience	1 = in the health science discipline, 0 = otherwise
science	1 = in the science discipline, 0 = otherwise
dadtertiary	a tertiary-educated father, 1 = yes, 0 = otherwise; a proxy of family socio-economic status
supportive	family supports of non-return intention 1 = supportive, 0 = otherwise
initialreturn	the initial intention prior to coming to New Zealand 1 = intend to return, 0 = otherwise
Perception-related variables	
goodHwage	perceived competitive wage at home country 1 = yes, 0 = otherwise
goodHoppk	perceived ample opportunities to apply specialized knowledge at home country, 1 = yes, 0 = otherwise
goodHlife	perceived good lifestyle at home country, 1 = yes, 0 = otherwise
goodHwenv	perceived favourable working environment at home country, 1 = yes, 0 = otherwise
goodHfren	perceived closer family and circle of friends at home country 1 = yes, 0 = otherwise
goodHrace	perceived equality of race at home country, 1 = yes, 0 = otherwise

Table 4: Binary logit model specification

Logistic regression			Number of obs =		516	
			LR chi2(23) =		239.12	
			Prob > chi2 =		0.0000	
Log likelihood = -236.22527			Pseudo R2 =		0.3360	
binaryreturn	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
q2_age	-.0146444	.1863991	-0.08	0.937	-.3799798	.3506911
agesq	-.0001005	.0032569	-0.03	0.975	-.0064838	.0062829
single	-.2695753	.4510498	-0.60	0.550	-1.153617	.6144661
male	-.0458707	.2391901	-0.19	0.848	-.5146746	.4229333
yrsinNZ	-.3081022	.175175	-1.76	0.079	-.651439	.0352345
yrsinNZsq	.028896	.0203665	1.42	0.156	-.0110216	.0688137
workyrs	-.1542194	.1489215	-1.04	0.300	-.4461001	.1376613
workyrssq	.0159128	.0115906	1.37	0.170	-.0068043	.03863
hselsewhr	-.2054634	.3229031	-0.64	0.525	-.8383418	.427415
phd	-.6088767	.349927	-1.74	0.082	-1.294721	.0769676
finalyear	.1811618	.2793177	0.65	0.517	-.3662908	.7286144
commerce	-.4048223	.3896004	-1.04	0.299	-1.168425	.3587804
hscience	-.8481386	.4001976	-2.12	0.034	-1.632512	-.0637657
science	-.3711961	.3560834	-1.04	0.297	-1.069107	.3267145
initialret~n	2.128113	.2662761	7.99	0.000	1.606221	2.650004
supportive	-.7250347	.2387555	-3.04	0.002	-1.192987	-.2570825
dadtertiary	.4174635	.2493822	1.67	0.094	-.0713166	.9062436
goodHwage	.5939854	.2728704	2.18	0.029	.0591692	1.128802
goodHwenv	.8170252	.3213904	2.54	0.011	.1871115	1.446939
goodHoppk	.7088585	.3045984	2.33	0.020	.1118566	1.30586
goodHlife	1.736497	.3061419	5.67	0.000	1.13647	2.336524
goodHfren	.9942648	.2805323	3.54	0.000	.4444316	1.544098
goodHrace	.1614864	.2646355	0.61	0.542	-.3571896	.6801624
_cons	-.3720289	2.600263	-0.14	0.886	-5.468451	4.724393

Table 5: Sample proportions between initial and current return intentions

Current intention	Initial intention		Total
	Not return	Return	
Not return	202 (85.6) (62.9)	34 (14.4) (17.4)	236 (100.0) (45.7)
Return	119 (42.5) (37.1)	161 (57.5) (82.6)	280 (100.0) (54.3)
Total	321 (62.2) (100.0)	195 (37.8) (100.0)	516 (100.0) (100.0)

Note:

Frequency
(Row percentage)
(Column percentage)

Table 6: Predicted probabilities of hypothesized scenarios

	Scenarios				
	1	2	3	4	5
q2_age	mean				
agesq	mean				
single	1				
male	1				
yrsinNZ*	mean			max	min
yrsinNZsq	mean			max	min
workyrs	mean			min	max
workyrssq	mean			min	max
hselsewhr	0	1			
phd*	0	1			
finalyear	0	1			
commerce	0				
hscience*	1				
science	0				
initialret~n*	1	0	0	0	0
supportive*	1				
dadtertiary*	1	0	0	0	0
goodHw~e*	1		0	0	0
goodHw~v*	1		0	0	0
goodHo~k*	1		0	0	0
goodHlife*	1		0	0	0
goodHfren*	1				
goodHrace	1		0	0	0
$\hat{\Pr}(return \mid X)$	0.9871	0.7611	0.0975	0.1827	0.6605
95% c.i.					
LB	0.9720	0.5046	0.0134	-0.1869	-0.1967
UB	1.0022	1.0176	0.1815	0.5524	1.5176

Note:
Scenario 1 is set as the baseline scenario, where subsequent scenarios diverge from.
LB: lower bound of the confidence interval
UB: upper bound of the confidence interval
* denotes significant variables.

Table 7: Change in predicted return probabilities

	from x = min	to x = max	Diff
q2_age	0.4689	0.3360	-0.1329
agesq	0.4499	0.3832	-0.0667
single	0.5086	0.4415	-0.0671
male	0.4528	0.4415	-0.0113
yrsinNZ	0.6419	0.0543	-0.5875
yrsinNZsq	0.3580	0.9558	0.5978
workyrs	0.4872	0.0559	-0.4313
workyrssq	0.4067	0.9917	0.5850
hselsewhr	0.4415	0.3916	-0.0499
phd	0.4415	0.3007	-0.1408
finalyear	0.4415	0.4865	0.0450
commerce	0.4415	0.3452	-0.0962
hscience	0.4415	0.2529	-0.1886
science	0.4415	0.3529	-0.0886
initialret~n	0.4415	0.8691	0.4276
supportive	0.4415	0.2768	-0.1646
dadtertiary	0.3424	0.4415	0.0991
goodHwage	0.4415	0.5887	0.1473
goodHwenv	0.4415	0.6415	0.2000
goodHoppk	0.4415	0.6162	0.1748
goodHlife	0.4415	0.8178	0.3763
goodHfren	0.2263	0.4415	0.2152
goodHrace	0.4415	0.4816	0.0401

Note:
For categorical variables, the minimum value the variables can take is zero while the maximum is one.

Table 8: Marginal effect on predicted probabilities

	dy/dx	Std. Err.	P>z	[95% C.I.]	
q2_age	-.0036109	.04596	0.937	-.09368	.08646
agesq	-.0000248	.0008	0.975	-.00159	.00154
single*	-.0671234	.11245	0.551	-.28752	.15328
male*	-.011339	.05907	0.848	-.12710	.10442
yrsinNZ	-.07597	.04249	0.074	-.15925	.00731
yrsinN~q	.007125	.005	0.154	-.00268	.01693
workyrs	-.0380265	.03668	0.300	-.10992	.03386
workyr~q	.0039237	.00287	0.171	-.00169	.00953
hselse~r*	-.0498867	.07814	0.523	-.20304	.10326
phd*	-.1407926	.07823	0.072	-.29412	.01254
finaly~r*	.0450242	.06933	0.516	-.09086	.18091
commerce*	-.096228	.09376	0.305	-.27999	.08754
hscience*	-.1885948	.09163	0.040	-.36819	-.0089
science*	-.0885878	.08639	0.305	-.25791	.08073
initia~n*	.4276179	.07108	0.000	.28829	.56694
suppor~e*	-.1646383	.05802	0.005	-.27835	-.0509
dadter~y*	.09908	.05884	0.092	-.01624	.21440
goodHw~e*	.1472768	.06647	0.027	.01699	.27755
goodHw~v*	.2000177	.07468	0.007	.05364	.34639
goodHo~k*	.1747798	.07267	0.016	.03235	.31720
goodHl~e*	.3762984	.06873	0.000	.24158	.51101
goodHf~n*	.2151944	.0641	0.001	.08955	.34083
goodHr~e*	.0401104	.06572	0.542	-.08869	.16892

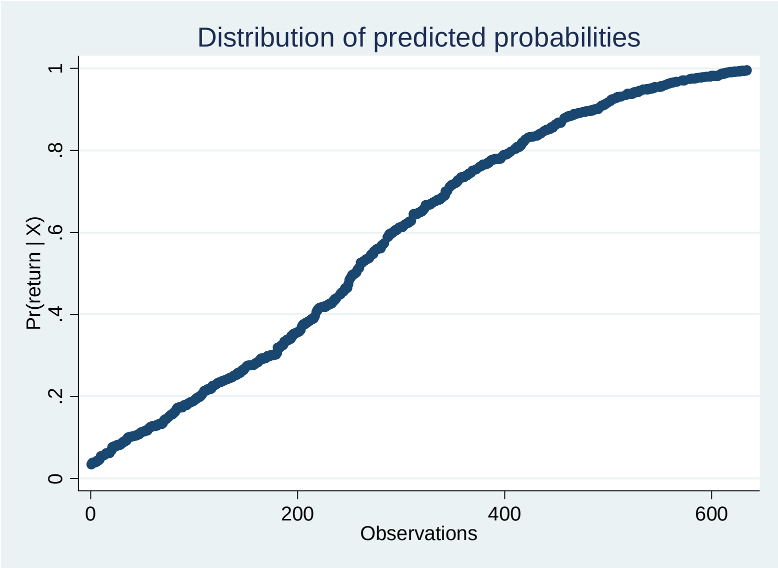
Note: (*) dy/dx is for discrete change of dummy variable from 0 to 1.

Table 9: Odds ratio (return vs not return)

	b	z	P>z	e^b	%
q2_age	-0.01464	-0.079	0.937	0.9855	-1.5
agesq	-0.00010	-0.031	0.975	0.9999	-0.0
single	-0.26958	-0.598	0.550	0.7637	-23.6
male	-0.04587	-0.192	0.848	0.9552	-4.5
yrsinNZ	-0.30810	-1.759	0.079	0.7348	-26.5
yrsinNZsq	0.02890	1.419	0.156	1.0293	2.9
workyrs	-0.15422	-1.036	0.300	0.8571	-14.3
workyrssq	0.01591	1.373	0.170	1.0160	1.6
hselsewhr	-0.20546	-0.636	0.525	0.8143	-18.6
phd	-0.60888	-1.740	0.082	0.5440	-45.6
finalyear	0.18116	0.649	0.517	1.1986	19.9
commerce	-0.40482	-1.039	0.299	0.6671	-33.3
hscience	-0.84814	-2.119	0.034	0.4282	-57.2
science	-0.37120	-1.042	0.297	0.6899	-31.0
initialret~n	2.12811	7.992	0.000	8.3990	739.9
supportive	-0.72503	-3.037	0.002	0.4843	-51.6
dadtertiary	0.41746	1.674	0.094	1.5181	51.8
goodHwage	0.59399	2.177	0.029	1.8112	81.1
goodHwenv	0.81703	2.542	0.011	2.2638	126.4
goodHoppk	0.70886	2.327	0.020	2.0317	103.2
goodHlife	1.73650	5.672	0.000	5.6774	467.7
goodHfren	0.99426	3.544	0.000	2.7027	170.3
goodHrace	0.16149	0.610	0.542	1.1753	17.5

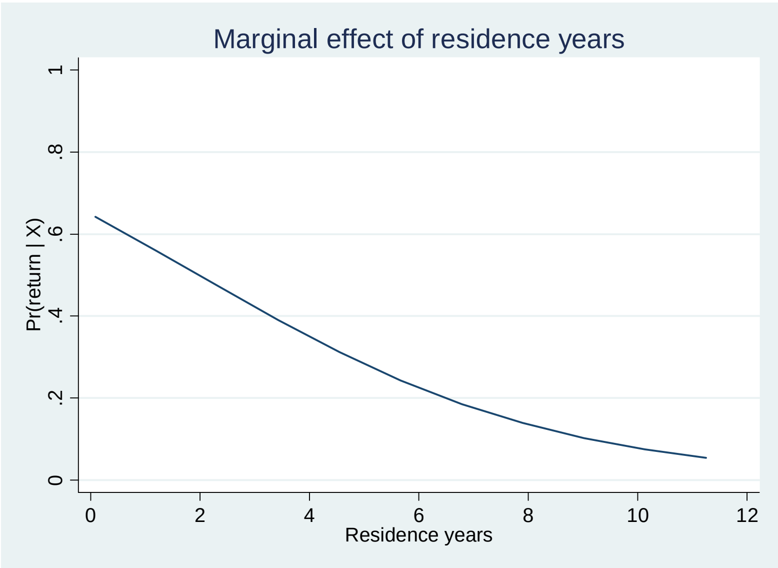
Note:
b = raw coefficient
z = z-score for test of b=0
P>z = p-value for z-test
e^b = exp(b) = factor change in odds for unit increase in X
% = percent change in odds for unit increase in X

Figure 1: Cumulative distribution function of the predicted probability



Note:
The sigmoid-shaped curve represents the cumulative distribution function of the predicted probabilities and approximates a logistic distribution.

Figure 2: Marginal effect of residence years on predicted return probabilities



References

Agarwal, V. B., & Winkler, D. R. (1984). Migration of Professional Manpower to the United States. *Southern Economic Journal, 50*(3), 814-830.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *The Journal of Political Economy, 70*(5), 9-49.

Becker, G. S. (1964). *Human Capital: A theoretical and empirical analysis with special reference to education*. London: University of Chicago Press.

Beine, M., Docquier, F., & Rapoport, H. (2001). Brain drain and economic growth: theory and evidence. *Journal of Development Economics, 64*(1), 275.

Beine, M., Docquier, F., & Rapoport, H. (2003). *Brain drain and LDC's growth: Winners and losers* (No. IZA DP No. 819). Bonn: Institute for the Study of Labor (IZA).

Ben-Akiva, M., & Lerman, S. R. (1989). *Discrete choice analysis: Theory and application to travel demand*. London: MIT Press.

Bhagwati, J., & Hamada, K. (1974). The brain drain, international integration of markets for professionals and unemployment. *Journal of Development Economics, 1*(1), 19.

Bhagwati, J., & Rodriguez, C. (1975). Welfare-theoretical analyses of the brain drain. *Journal of Development Economics, 2*(3), 195.

Bhagwati, J. N., & Dellalfar, W. (1973). The Brain Drain and Income Taxation. *World Development, 1*(1-2), 94-101.

Bratsberg, B. (1995). The Incidence of Non-return among Foreign Students in the United States. *Economics of Education Review, 14*(4), 373-384.

Brown, R. P. C., & Connell, J. (2004). The migration of doctors and nurses from South Pacific island nations. *Social Science & Medicine, 58*, 2193-2210.

Chen, T. J., & Su, H. Y. (1995). On-the-job training as a cause of brain drain. *Weltwirtschaftliches Archiv, 131*(3), 526-541.

Domencich, T. A., & McFadden, D. (1975). *Urban travel demand : a behavioral analysis* Amsterdam: North-Holland Publishing.

Gani, A., & Ward, B. D. (1995). Migration of professionals from Fiji to New Zealand: A reduced form supply-demand model. *World Development, 23*(9), 1633-1637.

Grubel, H. B., & Scott, A. D. (1966). The international flow of human capital. *American Economic Review, 56*(2), 268.

Gungor, N. D. (2003). *Brain drain from Turkey: an empirical investigation of the determinants of skilled migration and student non-return.* Unpublished Doctor of Philosophy, Middle East Technical University, Ankara.

Hamada, K., & Bhagwati, J. (1975). Domestic distortions, imperfect information and the brain drain. *Journal of Development Economics, 2*(3), 265.

Heckman, J. J. (2000). Microdata, heterogeneity and the evaluation of public policy. *Economic Sciences*, 254-322.

Huang, W.-C. (1988). An Empirical Analysis of Foreign Student Brain Drain to the United States. *Economics of Education Review, 7*(2), 231-243.

Johnson, H. G. (1967). Some economic aspects of the brain drain. *Pakistan Development Review, 7*, 379-411.

Kao, C. H. C., & Lee, J. W. (1973). An empirical analysis of China's brain drain into the United States. *Economic Development and Cultural Change, 21*(3), 500-513.

Kwok, V., & Leland, H. (1982). An economic model of the brain drain. *American Economic Review, 72*(1), 91.

Lee, E. S. (1966). A theory of migration. *Demography, 3*(1), 47-57.
Lien, D.-H. D. (1987). Economic analysis of brain drain. *Journal of Development Economics*, *25*(1), 33-43.

Lien, D.-H. D. (1988). Appropriate scientific research and brain drain: A simple model. *Journal of Development Economics, 29*(1), 77-87.

Lien, D. (2006). International accreditation and brain drain: A simple model. *Economics of Education Review, 25*(3), 335-340.

Lien, D. (2008). Quality assurance program and brain drain. *Education Economics, 16*(1), 59-73.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*: SAGE Publications, Inc.

Lowell, B. L., & Findlay, A. (2001). *Migration of highly skilled persons from developing countries: Impact and policy responses* (No. 44). Geneva: International Labour Office.

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision, 8*, 229-254.

Merwood, P. (2007). *International students: Studying and staying on in New Zealand* Department of Labour

Miyagiwa, K. (1991). Scale economies in education and the brain drain problem *International Economic Review, 32*(3), 743.

Mountford, A. (1997). Can a brain drain be good for growth in the source economy? *Journal of Development Economics, 53*(2), 287.

Myers, R. G. (1973). International Education, Emigration, and National Policy (A Longitudinal Case Study of Peruvians Trained in the United States). *Comparative Education Review, 17*(1), 71-90.

Schiff, M. (2005). Brain gain: claims about its size and impact on welfare and growth are greatly exaggerated: The World Bank, Policy Research Working Paper Series: 3708.

Sjaastad, L. A. (1962). The Costs and Returns of Human Migration. *The Journal of Political Economy, 70*(5), 80-93.

Stark, O., Helmenstein, C., & Prskawetz, A. (1997). A brain gain with a brain drain. *Economics Letters, 55*(2), 227-234.

Stark, O., Helmenstein, C., & Prskawetz, A. (1998). Human capital depletion, human capital formation, and migration: A blessing or a 'curse'? *Economics Letters, 60*(3), 363-367.

Stark, O., & Yong, W. (2002). Inducing human capital formation: migration as a substitute for subsidies. *Journal of Public Economics, 86*(1), 29.

Sue, V. M., & Ritter, L. A. (2007). *Conducting online surveys*. London: SAGE Publications.

Train, K. (1986). *Qualitative choice analysis: Theory, econometrics and an application to automobile demand*. London: MIT Press.

Train, K. E. (2003). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press.

Vidal, J.-P. (1998). The effect of emigration on human capital formation. *Journal of Population Economics, 11*(4), 589-600.

Vinokur, A. (2006). Brain migration revisited. *Globalisation, Societies & Education, 4*(1), 7-24.

Waldorf, B. (1995). Determinants of international return migration intentions. *Professional Geographer, 47*(2), 125.

Winkelmann, R., & Boes, S. (2006). *Analysis of microdata*. New York: Springer.

Wong, K.-Y. (1995). *International trade in goods and factor mobility*. London: MIT Press.

Zweig, D. (1997). To return or not to return? Politics vs. economics in. *Studies in Comparative International Development, 32*(1), 92.