

CAPÍTULO V: RESULTADOS

5.1. Introducción

Una vez resumida la metodología y hecho el análisis descriptivo de la base de datos, en este capítulo se plantean los resultados a los que se llega, sobre la probabilidad que tiene un remigrante mexicano de abrir un negocio.

La secuencia de este apartado es la siguiente: en primer lugar se hace una comparación entre las tablas estadísticas con las observaciones originales y con las observaciones finales de las regresiones para evaluar el problema de autoselección; después se muestran los resultados que se obtienen en los dos modelos utilizando el modelo probit; y por último, se realiza la instrumentación de las variables de ahorro.

5.2. Comparación de las Tablas Estadísticas

Comparando las dos tablas estadísticas de la muestra original y de la muestra utilizada, no se observa una diferencia importante entre las variables dependientes de las dos muestras. Se puede decir que cerca del 14% de los individuos invierten en la formación de un negocio fijo, mientras que sólo el 2% lo hace en un negocio agrícola. Otras variables que muestran pequeñas diferencias son las de *número de hijos*, *número de viajes*, *ahorros traídos a México* y *ahorros en los Estados Unidos*. Ver Tabla 14.

En cuanto a la *edad*, se observa que la edad promedio de los migrantes es relativamente menor en la muestra utilizada que en la muestra original. Con relación a la variable *edad de retorno del primero y último viaje*, se advierte que la muestra utilizada

representa a individuos que retornan más jóvenes a México. Por el contrario, en cuanto a la *duración de la permanencia en los Estados Unidos*, la muestra ocupada representa a individuos que en promedio permanecen más tiempo en el país huésped.

La muestra con las 3563 observaciones, representa a individuos con mayor *educación*, que tienen un mayor *conocimiento del inglés*, y que son encuestadas en comunidades *urbanas*, considerando la clasificación antes descrita. En cuanto a la *ocupación realizada en los Estados Unidos*, esta muestra sobrerrepresenta a las personas que trabajan en la industria, que son *peones en la industria, a los comerciantes y a los que trabajan en establecimientos*. Por otra parte, las ocupaciones de *operador de máquina y trabajador de vigilancia* quedan subrepresentadas levemente en la muestra.

En lo que refiere al *lugar de residencia en los Estados Unidos*, la muestra utilizada es más representativa de los estados de *Nueva York, Pennsylvania y Idaho*, y menos representativa de los estados de *Nevada y Colorado*. Las diferencias sin embargo, son mínimas. Con relación a los *años de encuesta*, la muestra ocupada en la regresión tiende a sobrerrepresentar a los individuos encuestados de *1992 a 1994*, y después de *1997 al 2000*. Los años antes de *1990* quedan claramente subrepresentados en esta muestra.

Se observa que la pérdida de observaciones se debe en su mayoría a las variables de *ahorros en los Estados Unidos* (1118)¹, *ahorros traídos a México* (932), *remesas* (511), *conocimiento del idioma* (140), *ocupación en los Estados Unidos* (84), y *lugar de residencia en los Estados Unidos* (27).

¹ El número entre paréntesis equivale al número de observaciones faltantes de cada una de las variables

Dado que las variables *ahorros en los Estados Unidos*, *ahorros traídos a México* y *remesas* son las variables con el mayor número de observaciones faltantes, excluyendo estas variables del modelo se ganan cerca de mil quinientas observaciones. A fin de obtener las implicaciones de la pérdida de observaciones, en cada modelo se llevan a cabo dos regresiones: la primera, incluyendo todas las variables del modelo; y la segunda, dejando de lado las tres variables causantes de una pérdida considerable de observaciones.

Lo que se busca con la exclusión de estas variables es poder comparar las dos regresiones y ver si existen grandes diferencias entre ellas. De esta forma, se puede concluir si la pérdida de observaciones conlleva a resultados sesgados o no. A continuación se muestran los resultados empleando el modelo probit.

5.3 Resultados del Modelo Probit

5.3.1. Modelo 1: Negocios con Establecimiento

En este modelo como en el siguiente, la variable dependiente se define como la decisión de abrir un negocio después del primer viaje migratorio. Se deja fuera de estudio a aquellos negocios que se tienen antes de migrar. Esto, con el objetivo de medir exclusivamente el impacto que la acumulación de capital en el exterior tiene sobre la decisión de abrir un negocio. En este modelo la variable dependiente es *negfijo*.

En la Tabla 15. y Tabla 16. se muestran los resultados obtenidos para las dos regresiones del modelo. Como se explica en párrafos anteriores, la única diferencia entre las dos regresiones radica en la exclusión de las variables *ahorros en los Estados Unidos*,

ahorro traído a México y remesas de la segunda ecuación. En general, los resultados de ambas ecuaciones son bastante similares entre si.

En ambos casos se observa que la variable *edad* tiene un coeficiente positivo y significativo al 5%. El coeficiente de la primera regresión puede interpretarse de la siguiente manera: al cumplir el jefe de familia un año más de vida, incrementa la probabilidad de que invierta en un negocio en un 0,45%. Esto quiere decir que, cuanto mayor es el jefe de familia, mayor es la probabilidad de que invierta en este rubro. Este resultado converge con la idea de que a medida que aumenta la edad del individuo, aumenta su interés por crear una fuente de ingreso segura para el momento de su retiro del mercado laboral.

Con relación al *estado civil* del jefe de familia, se encuentra que la variable categórica *soltero* es negativa y significativa al 5% en ambas regresiones. Esto supone una relación negativa entre convertirse en autoempleado y ser soltero. A medida que un jefe de familia continúa soltero a su retorno de la experiencia migratoria, disminuye su probabilidad de abrir un negocio en un 8,8% con relación al jefe de familia casado.

De igual forma, la variable categórica *es papá*, es negativa y significativa al 10% en la primer regresión, y al 5% en la segunda regresión. Estos coeficientes señalan que la probabilidad de que un individuo con hijos abra un negocio a su retorno disminuye en un 5,3% con relación a los que no tienen hijos.

Tanto para la variable *soltero*, como para la variable *es papá*, se puede usar la misma justificación: un individuo soltero o sin hijos tiene una mayor propensión a aceptar trabajos menos seguros, o que requieren de una mayor movilidad en comparación a los individuos casados y con hijos. Esto se debe a que los solteros y las personas sin hijos tienen menores responsabilidades y pueden dedicarse de lleno a un trabajo que es mejor remunerado en el mercado laboral.

Para el caso del nivel de educación de los jefes de familia, los coeficientes de las variables de nivel de *educación primaria* y *secundaria completa*, junto con el nivel de *educación superior incompleta* salen significativos al 5% en ambas regresiones. En la segunda regresión sin embargo, el coeficiente de la variable *sin escolaridad* es negativo y significativo al 10%, y con la misma significancia pero con signo positivo el de la variable *secundaria y preparatoria incompleta*.

La discrepancia en estos resultados puede deberse al problema de pérdida de observaciones en las regresiones. En la primer regresión, donde la variable *sin escolaridad* no es significativa, se trabaja con una muestra que sobrerrepresenta a aquellos jefes de familia con un nivel de educación mayor y subrepresenta a aquellos individuos sin escolaridad. Esto, con relación a la muestra original.

A pesar de estas diferencias, los resultados obtenidos no se contradicen entre si. En la primer regresión se observa que la probabilidad de invertir en un negocio aumenta en un 5,3% si se pasara de primara incompleta a primaria completa. En relación a la secundaria completa y universidad incompleta, esta probabilidad aumenta en un 9,2% y un 8,8%

respectivamente con relación a la primaria incompleta. En la segunda regresión, la probabilidad de invertir en un negocio disminuye en un 2.9%, de pasar de primaria incompleta a sin escolaridad.

A medida que el individuo concluye su estudio elemental y medio, aumenta la probabilidad de que éste a su retorno invierta en un negocio. Un individuo analfabeta es menos probable que designe recursos hacia ese rubro. El nivel de educación por lo tanto, si puede considerarse como un factor decisivo para la apertura de un establecimiento.

Por otra parte, únicamente en la segunda regresión, el coeficiente del *número de hijos* es positivo y significativo al 5%. Es decir, a medida que aumenta el número de hijos en el hogar en una unidad, aumenta la probabilidad de que se abra un negocio en un 0,50%.

La explicación a este resultado es la contraparte de la variable *soltero y es papá*. A medida que aumenta el número de hijos, aumenta la necesidad de obtener fuentes de ingresos seguros. Y además, a medida que aumenta el número de hijos, aumenta el número de personas que puede cuidar del negocio. De hecho, como se obtiene en la descripción estadística de los negocios, en estos establecimientos la mayoría de los trabajadores son familiares.

En el caso de la *edad de retorno del primer viaje*, en ambas regresiones se observa un impacto negativo de la variable sobre la probabilidad de formar un negocio. En la primer regresión se advierte que al incrementarse la edad del migrante, que retorna en un año, disminuye la probabilidad de que se invierta en un establecimiento en un 0,29%.

La justificación es la siguiente: cuanto más grande regresa el individuo de su primer viaje, mayores son las responsabilidades a las que se enfrenta el individuo a su regreso. Por ejemplo, la edad de retorno promedio del primer viaje es de veintiocho años. A esta edad, el individuo generalmente busca incorporarse al mercado laboral y destina gran parte de sus ingresos en la manutención de la familia y en la compra de bienes no duraderos.

El *conocimiento del inglés que se tiene*, parece ser un factor importante en la decisión de convertirse en autoempleado. Tanto en la primer como segunda regresión, la variable que denota una buena comprensión del idioma es positiva y significativa al 10 y al 5%, respectivamente. Con relación a la primera regresión, el coeficiente muestra que el tener un buen conocimiento de la lengua aumenta la probabilidad de abrir un negocio en un 4,8%, con respecto al que no habla y no entiende.

La *ocupación en los Estados Unidos* también es importante en la decisión de convertirse en autoempleado. En ambas regresiones se cumple que el coeficiente de las variables *trabajador en una industria, operador de máquina, peón en una industria, trabajador en un establecimiento y comerciante* es positivo y significativo al 5%.

Esto conduce a suponer que los individuos durante su permanencia en el exterior atraviesan por un proceso de acumulación de capital humano. Y una vez adquirido ciertos conocimientos y capacitaciones, deciden poner en práctica lo aprendido abriendo un negocio.

En la primer regresión, en la cual se incluyen las variables relacionadas con la acumulación de capital financiero, se obtiene que los coeficientes de la variable de *ahorros traídos a México* y la de *remesas* son positivos y significativos al 10 y al 5% respectivamente. Este resultado se puede interpretar de la siguiente manera: el aumentar la cantidad de dólares traídos a México en mil dólares, aumenta la probabilidad de que un remigrante invierta en un negocio fijo en un 0,08%. Y el aumentar la cantidad remitida a México mensualmente en mil dólares, aumenta la probabilidad en un 2,6%.

Estos resultados demuestran que la acumulación de capital financiero juega un papel importante en la decisión de los individuos de convertirse en autoempleados. Una de las causas por las que esto es así, como lo señala Mesnard (1999) es la existencia de restricciones al capital en el país de origen, lo que provoca la necesidad en el individuo de migrar a fin de hacerse de formas de financiamiento alternas.

En cuanto al *lugar de encuesta*, los resultados de ambas regresiones muestran que el coeficiente de la variable categórica *pueblo* es positivo y significativo al 5%. La probabilidad de abrir un negocio en un pueblo es 3,8% mayor que de abrirlo en un rancho. El *lugar de residencia en los Estados Unidos* también parece tener un impacto en la decisión de abrir un negocio. El haber estado en Nueva York durante el último viaje por ejemplo, aumenta la probabilidad de convertirse en autoempleado en un 20,8% en comparación a cualquier otro lugar diferente a los especificados en el modelo. En el caso de Idaho, la probabilidad aumenta en un 18,01%.

Por último, los *años de encuesta* en ambas ecuaciones señalan que en los años 1982 y 1998, en contraparte a los años 1990 y 1987 aumentaron las dificultades de abrir un negocio. En 1998 por ejemplo, la probabilidad de abrir un negocio disminuye en 4,6% con relación al año de referencia 1991. Para el caso del año 1982, la probabilidad disminuye en 7,3%. Por el contrario, en el año 1987, la probabilidad aumenta en casi 15%, y en 1990, en 7,69% con respecto a 1991.

Según Cárdenas (1999) los años de 1981 a 1988 corresponden a un periodo caracterizado por la crisis de la deuda y el estancamiento económico, mientras que de 1988 a 1994 predomina la estabilización económica y los cambios estructurales en el país. La situación económica durante el año de referencia 1991, es de recuperación y cambios.

Se observa, que durante los años con mejoras en la economía del país aumenta la probabilidad de abrir un negocio, mientras que en los años de crisis e incertidumbre económica, disminuye tal probabilidad. Lo anterior muestra que la situación macroeconómica del país también repercute sobre la decisión de convertirse en autoempleado

5.3.2. Modelo 2: Negocio agrícola

A diferencia del otro modelo como se menciona en el Capítulo III, en este caso la variable de *negagricultura* abarca únicamente la inversión realizada en la agricultura y en la crianza de ganado. Los resultados obtenidos, se muestran a continuación. Ver Tabla 17 y Tabla 18.

La *edad* del jefe de familia en este modelo también juega un papel importante en la decisión de invertir en un negocio agrícola. El coeficiente de esta variable es positivo y significativo al 5%. En la primera ecuación se observa que al cumplir el jefe de familia un año más de vida, se incrementa la probabilidad de que éste invierta en un negocio agrícola en un 0,03%.

La causa de esto puede ser la misma que la del modelo anterior, la búsqueda por afianzar una fuente de ingreso alterna. O simplemente puede deberse, a diferencia del modelo anterior, a que los jefes de familia se dedican a las tareas agrícolas por herencia. Desafortunadamente, esta última hipótesis es tan solo una suposición, la falta de variables no permite la comprobación de tal argumento. Cabe mencionar que en este modelo a comparación del primero, el cambio en la probabilidad de abrir un negocio dado el aumento en la edad es menor.

El nivel de *educación* de los jefes de familia, y el *estado civil* de los mismos parece no repercutir significativamente en la decisión de abrir un negocio agrícola. Por otra parte, el coeficiente de la variable *edad de retorno del primer viaje*, igual que en el modelo anterior, es negativo y significativo al 5%. En la primera regresión se advierte que al incrementarse la edad del migrante que retorna en un año, disminuye la probabilidad de que se invierta en un negocio agrícola en 0,06%. El argumento del modelo anterior es válido también para este modelo.

En cuanto al impacto que la acumulación de capital tanto financiero como humano tienen en la decisión de invertir en un negocio agrícola, se encuentra evidencia en la primer

regresión, en oposición a la segunda, del impacto que esta acumulación tiene. El coeficiente de la variable *ahorro en los Estados Unidos* es positivo y significativo al 5%, al igual que el de la variable *jefe de industria, trabajador en industria y peón en industria*.

Con estos resultados se puede sospechar, que los ahorros junto con la capacitación en el sector industrial durante la estancia en los Estados Unidos constituyen elementos que estimulan la inversión en este rubro. Sin embargo, no es clara esta relación debido a que en la segunda regresión no se presentan los mismos resultados.

En relación al vínculo existente entre el *tipo de tierra* que poseen los migrantes en el momento de la encuesta y la probabilidad de que estos inviertan en un negocio agrícola, se observa que los migrantes con pastizales aumentan su probabilidad de abrir un negocio con respecto a otro tipo de tierra en 2,94%. Por otro lado, en relación a la *tenencia de la tierra*, se observa que los hogares con tierras propias aumentan su probabilidad de abrir un negocio agrícola en 2,5% con relación a los que arriendan la tierra.

A diferencia del modelo pasado, en este modelo no queda claro el papel que juega el *lugar de encuesta* en la decisión del jefe de familia de invertir en este tipo de negocios. En la primer regresión se observa que el coeficiente de la variable *pueblo* es negativo y significativo al 5%. Al pasar el lugar de encuesta de rancho a pueblo, disminuye la probabilidad de convertirse en autoempleado en 0,70%. En la segunda regresión sin embargo, se tiene que el coeficiente de la variable *areametropolitana* es positivo y significativo al 5%. Es decir, al pasar de rancho a un área metropolitana aumenta la probabilidad de invertir en un negocio en 1,86%.

A pesar de que ambas regresiones no comparten las mismas variables significativas, los coeficientes si tienen los mismos signos y los efectos no se contradicen entre si. El resultado obtenido sugiere que en comparación con el rancho, aumenta la probabilidad de abrir un negocio agrícola en áreas metropolitanas, y a su vez disminuye la probabilidad de abrirlo en un pueblo.

Este resultado es consistente con las características presentadas sobre la base de datos. En la región oeste del país, a diferencia de la región centro y sur, la actividad agrícola no se limita a los pequeños pueblos o ranchos, sino que la actividad agrícola se extiende también hacia las áreas urbanas y sus alrededores. Este resultado es congruente con la región de estudio de esta base.

En cuanto al *lugar de residencia en los Estados Unidos*, el coeficiente de la variable *florida* se muestra positivo y significativo al 5%. El vivir en Florida en vez de cualquier otro estado no enumerado en el modelo aumenta la probabilidad de convertirse en autoempleado en 4,6%. Este resultado muestra que el vivir en un estado agrícola en el país huésped, favorece la inversión en un negocio agrícola en el país de origen.

A diferencia de la inversión en un negocio fijo, la inversión en un negocio agrícola muestra haber tenido un desempeño no favorable en los últimos años. De 1992 al 2002 parece haberse agudizado esta situación. Las cifras, con relación al año base *19991*, muestran que a medida que aumenta el año de encuesta, disminuye la probabilidad de que se invierta en un negocio agrícola. Estos resultados son el reflejo de la situación del campo mexicano.

5.4. Instrumentación de las Variables Endógenas

En varios estudios realizados, se ha encontrado que la decisión de ahorrar está determinada endógenamente en la ecuación de decisión de invertir en un negocio. Ilahi (1999) plantea que aquellos individuos con interés de convertirse en autoempleados a su retorno, ahorran más durante su permanencia en el exterior que los que no. Como se ha discutido en el Capítulo III, la existencia de variables endógenas dentro del modelo causa sesgo en los resultados. Para tratar este problema, se busca la instrumentación de las variables de ahorro.

El modelo cuenta con tres variables relacionadas con la acumulación de capital financiero en los Estados Unidos, que son: *ahorros acumulados mensualmente en los Estados Unidos*, *ahorros traídos a México al retorno del migrante*, y *remesas enviadas mensualmente a México durante la permanencia del migrante en el exterior*. Dado que la correlación entre las tres variables resulta ser bastante baja, tratamos la endogeneidad de cada una de las variables por separado. Ver Tabla 30.

5.4.1. Modelo 1: Negocio con Establecimiento

En el trabajo se corren tres regresiones distintas para cada modelo, en cada regresión se incluye únicamente una de las variables de ahorro. Después se instrumenta cada una de esas variables con diferentes variables instrumentales. El resultado obtenido se presenta a continuación.

5.4.1.1. Caso 1: Ahorros acumulados mensualmente en los Estados Unidos.

Al correr el modelo de negocio fijo utilizando únicamente la variable *ahorros en los Estados Unidos*, el número de observaciones asciende a 4041, que a diferencia del modelo

original, el número corresponde a 3575. Los resultados entre ambas regresiones son bastante similares con las siguientes diferencias.

En esta regresión, el coeficiente de la variable *sinescolaridad* no sólo es negativo, sino que también es significativo a un 10%. De igual forma, el coeficiente de la variable *número de hijos* también adquiere significancia del 10%. La variable relacionada con el buen conocimiento del inglés se muestra significativa al 5%, lo cual reafirma la idea de que una mayor acumulación de capital humano conlleva a una mayor probabilidad de convertirse en autoempleado. En cuanto al tipo de comunidad encuestada, el coeficiente de la variable *urbanopequeño* también aparece significativo al 10%, Ver Tabla 19.

En el caso de la variable *ahorros en los Estados Unidos*, ésta es instrumentada con las variables *ingreso por mes durante el primer y último viaje*, *gasto en comida por mes* y *gasto en renta por mes en los Estados Unidos*. La variable *cuenta bancaria*, a pesar de presentar una correlación con la variable endógena se excluye de la instrumentación, debido a la posible correlación de ésta con la variable dependiente del modelo negocio fijo y por no aportar grandes diferencias al modelo cuando se incluye. Ver Tabla 25 y Tabla 28.

Una vez hecha la instrumentación del *ahorro en los Estados Unidos* con las variables mencionadas en el párrafo anterior, queda corregido el sesgo ocasionado por la endogeneidad de la variable, y se observa que no existen grandes diferencias entre las dos regresiones. Se concluye que los resultados presentados en el modelo original son consistentes. Ver Tabla 20.

5.4.1.2. Caso 2: Ahorros traídos a México

En este caso, el número de observaciones en la regresión aumenta a 4230. Y los resultados obtenidos en esta regresión son similares a los del modelo original. Las diferencias que se presentan en este caso son las siguientes: el coeficiente de la variable *número de hijos* sale positivo y significativo al 10%; el coeficiente de la variable *hablaentiendemucho* inglés resulta también significativa pero al 5%; y finalmente el coeficiente de la variable *trabajador en educación* es significativo al 10%. Ver Tabla 19.

El escoger una variable instrumental para la variable *ahorros traídos a México* resulta ser una tarea bastante difícil dado que las correlaciones entre las variables instrumentales y la variable endógena son muy pequeñas. Se llevan a cabo dos instrumentaciones para este caso. En principio se adoptan como variables instrumentales, las variables con correlación positiva con la variable endógena y sin correlación con la variable dependiente del modelo. Las variables seleccionadas son *ingreso por mes durante el primero y último viaje*, *gasto en comida por mes* y *meses trabajados por año en los Estados Unidos*. Ver Tabla 26 y Tabla 28.

Instrumentando con esas variables, obtenemos los mismos resultados que los hallados en el modelo original. La única diferencia importante es que la variable *ahorros traídos a México* pierde su significancia. Sin instrumentar, el coeficiente de la variable es de 0,0047. Una vez realizada la instrumentación el coeficiente de la misma aumenta a 0,0051. Debido a este aumento en el coeficiente de la variable, se concluye que la endogeneidad de esta variable causa que la variable esté subestimada. Ver Tabla 21.

Para comprobar la veracidad de ese resultado, se lleva a cabo una segunda instrumentación, pero ahora adicionando a la primera instrumentación las variables *gasto en renta por mes*, *número de amigos* y *horas trabajadas por semana*. Las primeras dos variables se correlacionan negativamente con la variable endógena y la tercera variable parece tener cierta correlación con la variable dependiente del modelo, mas se incluye por tener una correlación relativamente alta con la variable instrumentada.

Los resultados obtenidos señalan ciertas diferencias con respecto a la regresión sin instrumentar. Esto específicamente con relación al estado civil del jefe de familia, duración del primer viaje, y al año de encuesta. Sin embargo, bajo esta instrumentación la variable *ahorros traídos a México* resulta significativa al 10%. Con la instrumentación, la variable adquiere nuevamente un coeficiente mayor, 0,0187. Por lo que en esta instrumentación también se concluye que la endogeneidad de la variable provoca que la variable esté subestimada.

Es difícil predecir la magnitud del sesgo, pero en base a las dos instrumentaciones realizadas se puede concluir que el sesgo es negativo, y que éste oscila entre 0,0004 y 0,0138. Ver Tabla 20.

5.4.1.3. Caso 3: Remesas

Incorporando únicamente la variable *remesas*, incrementa el número de observaciones en la regresión a 4656. En general los resultados son consistentes con los del modelo original. En este caso sin embargo, el coeficiente de la variable *uniónlibre* además de salir negativo sale significativo al 5%. Además, el coeficiente de *número de hijos* es

significativa pero al 10%. El nivel de conocimiento del inglés es positivo y significativo para las variables *nohablaentiendemucho* y *hablaentiendemucho*. Por último, el coeficiente de la variable *urbanopequeño* sale positivo y significativo al 5%. Ver Tabla 19.

En la instrumentación de esta variable, también se sigue la misma dinámica empleada con la variable anterior. La primera instrumentación se realiza ocupando variables con correlación tanto positiva como negativa, pero que no se correlacionan con la variable dependiente del modelo. Las variables instrumentales son *ingreso por mes durante el primero y último viaje*, *gasto en comida por mes*, *gasto en renta por mes* y *número de amigos en Estados Unidos*. Ver Tabla 27 y Tabla 28.

En esta instrumentación se obtiene que a diferencia del modelo original, las variables *es padre* y *remesas* pierden significancia, y a su vez las variables *unionlibre* y *areametropolitana* la ganan. En el caso de la variable *remesas*, con la instrumentación la variable pierde significancia. En la regresión original, el coeficiente de la variable es de 0,1272, mientras que en la regresión instrumentada el coeficiente disminuye a 0,1018. Por lo tanto se concluye que el sesgo causado por la endogeneidad de la variable sobrestima el efecto de la variable. Ver Tabla 21.

La segunda instrumentación incorpora como variables instrumentales únicamente los *ingresos percibidos mensualmente durante el primero y último viaje*, y además la variable *horas trabajadas por mes*. Nuevamente, el problema con esta última variable es su correlación con la variable dependiente del modelo, sin embargo se utiliza en la instrumentación debido a que presenta la mayor correlación con la variable *remesas*.

A diferencia de la regresión del modelo sin instrumentar, en este caso la variable *unionlibre*, al igual que la variable *número de hijos*, y *técnico* resultan significativas al 10%. En el caso de la variable *areametropolitana* la significancia aumenta al 5%. En esta regresión, a diferencia de la regresión anterior, la variable *remesas* sale positiva y significativa al 5%.

En este caso, el coeficiente de la variable *remesas* es de 0,2604, por lo que la endogeneidad de la variable ocasiona que la variable esté subestimada. Dado que en las dos instrumentaciones no se llega a un consenso sobre la dirección del sesgo, sólo se puede decir que la magnitud del sesgo oscila entre .0252 y .1332. Ver Tabla 20.

5.4.2. Modelo 2: Negocio Agrícola

5.4.2.1. Caso 1: Ahorros acumulados en los Estados Unidos

Al incorporar exclusivamente la variable *ahorros en los Estados Unidos* se ganan cerca de 486 observaciones con relación al modelo original. También varias variables adquieren significancia, como la variable *unionlibre* y *areametropolitana* del 5%, y la de *secundaria completa* y *duración del último viaje* del 10%. Fuera de eso, ambas regresiones son muy similares entre si. Ver Tabla 22.

Para esta variable se llevan a cabo dos instrumentaciones. La primera instrumentación se realiza utilizando únicamente las variables que tienen una mayor correlación con la variable endógena, que son *el ingreso por mes durante el primer y último viaje*, *el gasto en renta* y *la existencia de una cuenta bancaria en los Estados Unidos*. Todas las variables no presentan una correlación con la variable dependiente del modelo.

Al realizar la instrumentación se obtiene un estadístico F donde se acepta la hipótesis nula de que todas las variables son iguales a cero. Por lo tanto, se dejan de lado los resultados obtenidos en esta instrumentación. Ver Tabla 24.

En la segunda instrumentación, se incluyen como variables instrumentales a todas las variables correlacionadas con la variable *ahorro en los Estados Unidos* y no correlacionadas con la variable dependiente del modelo. Los resultados obtenidos bajo esta instrumentación son relativamente diferentes a los obtenidos sin instrumentación. A diferencia de la regresión sin instrumentación, en la regresión salen tanto la variable *divorciado* como la variable *separado* significativas al 5%. La variable *hablaentendealgo* de inglés junto con la variable *trabajador en educación* salen negativas y significativas al 10%. Y las variables *privado*, *pastizal* y *pueblo* dejan de tener significancia. Ver Tabla 23.

Con relación a la variable instrumentada, en ambas regresiones la variable de *ahorros en los Estados Unidos* resulta significativa y positiva. En la regresión sin instrumentación sin embargo, la significancia de la variable es del 5%, mientras que en la regresión instrumentada, ésta es del 10%. El coeficiente de la variable en la regresión original es de 0,1630, mientras que con la instrumentación el coeficiente disminuye a 0,0172. Por lo que en relación al sesgo, se concluye que la endogeneidad de la variable ocasiona una sobreestimación, y el sesgo generado es positivo. La magnitud del sesgo encontrado es de cerca de .1458.

5.4.2.2. Caso 2: Ahorros traídos a México

Utilizando la variable *ahorros traídos a México*, el número de observaciones en la regresión asciende a 3599, a diferencia de las 2930 de la del modelo original. Las diferencias más marcadas en ambas regresiones radica en la significancia de la variable *unionlibre* y *areametropolitana*. Ver Tabla 22.

Para corregir el problema de endogeneidad de esta variable, también se llevan a cabo dos instrumentaciones. Sin embargo, la instrumentación con las variables que se correlacionan positivamente con la variable endógena *ingreso por mes durante el primero y último viaje, meses trabajados al año y la existencia de cuenta bancaria*, conducen a un F estadístico cercano a uno, por lo que se omiten los resultados obtenidos bajo esa instrumentación. Ver Tabla 24.

En la segunda instrumentación se utilizan como variables instrumentales todas las variables correlacionadas con la variable endógena *ahorros traídos a México* tengan correlaciones positivos o negativos, excluyendo las variables correlacionadas con la variable dependiente del modelo. Estas variables son *ingreso por mes durante el primero y último viaje, meses trabajados por año, gasto en renta, número de amigos y la existencia de una cuenta bancaria*.

Los resultados obtenidos bajo esta instrumentación también difieren un poco de los resultados sin instrumentación. Utilizando variables instrumentales, se obtiene que la variable *divorciado* y *separado* son significativas al 10 y 5% respectivamente. También la variable de *preparatoria incompleta* aparece negativa y significativa al 10%. Las variables

privado y *pastizal* pierden significancia con la instrumentación. Tanto con como sin instrumentación, la variable *ahorros traídos a México* sale no significativa. Ver Tabla 23.

5.4.2.3. Caso 3: Remesas

En esta regresión, el número de observaciones es de 4095. Los valores obtenidos en las variables *unionlibre*, *nohablaentiendealgo*, *hablaentiendemucho* y *areametropolitana* difieren de aquellos del modelo original. La variable *unionlibre* adquiere significancia del 5%, mientras que las variables *nohablaentiendealgo* y *hablaentiendemucho* aumentan su significancia al 10%. Lo mismo sucede con la variable *areametropolitana*. Ver Tabla 22.

Con relación a la instrumentación, la variable *remesas* también presenta los mismos problemas que las dos variables anteriores. Al realizarse la instrumentación con las variables *ingreso por mes durante el primero y último viaje* y *horas trabajadas a la semana en los Estados Unidos*, se obtiene un F estadístico cercano a uno. Por lo que no se considera el modelo. Ver Tabla 23.

La otra instrumentación utiliza como variables instrumentales las mismas usadas en el caso anterior, sólo que *horas*, es sustituida por las variables *gasto en renta*, *número de amigos* y *cuenta bancaria*. A diferencia de la variable *horas*, estas variables no se correlacionan con la variable dependiente. Bajo esta instrumentación, los resultados obtenidos también difieren sustancialmente de los resultados de la regresión sin instrumentación. La variable de *separado*, junto con la de *conductor* salen negativas y significativas al 5 y 10%, respectivamente.

En cuanto a la variable *remesas*, a diferencia de la regresión sin instrumentación, bajo instrumentación el coeficiente de la variable sale positivo y significativo al 10%. El coeficiente de la variable en la regresión sin instrumentar es de -0,0417, mientras que bajo la instrumentación, el coeficiente asciende a 0,0637. En base a estos resultados, se concluye que la endogeneidad de la variable genera una subestimación muy grande de la variable. Sin embargo, dado que los signos de los coeficientes en ambas regresiones difieren entre si, existen dudas sobre la validez de las variables instrumentales para resolver el problema de endogeneidad. Ver Tabla 24.