

Impacto del Internet en el Ingreso Laboral: Análisis del lugar y manejo del Internet^{*}

Diego Martin^{**}

Director: Juan Fernando Vargas

Junio, 2015

Resumen

Este estudio analiza el efecto del uso de Internet sobre el ingreso laboral, eliminado el problema de endogeneidad que existe por motivos de selección. También se analizan los efectos según el lugar de acceso y las actividades realizadas en Internet. Para esto se utiliza Propensity Score Matching, tomando información para Colombia entre el año 2009 a 2011. Los resultados muestran que existe un efecto positivo, significativo y heterogéneo entre trabajadores asalariados y cuenta propia, siendo los últimos los más beneficiados. Con respecto al lugar de uso se encuentra un impacto positivo del uso del Internet en el trabajo y/o en el hogar, con una mayor magnitud en el primer sitio de acceso. Para el tema del manejo del Internet, no se encuentra efecto cuando se utiliza en actividades de entretenimiento, pero si cuando se usa en tareas educativas diariamente. Esto sugiere que usar este servicio aumenta la productividad y el capital humano, lo que se transforma en un incremento en los ingresos. Por tanto, se encuentra que el lugar de uso es una señal informativa para el mercado laboral, mientras que el impacto del manejo de esta herramienta depende de la frecuencia con que se realicen las actividades.

^{*}Documento para tesis de grado en maestría de la Universidad del Rosario.

^{**}Asistente de Investigación y estudiante de Maestría en economía en la Universidad del Rosario. Estoy muy agradecido con Juan F. Vargas por su dedicación y buen asesoramiento como director de tesis. Además agradezco a Joaquin Coleff, Juan Carlos Guataquí y al seminario de estudiantes del departamento de económica de la Universidad del Rosario (SETE) por sus valiosos comentarios. Así como a los profesores del curso de Economía Laboral Juan Miguel Gallego y Carlos Sepúlveda por ayudarme a empezar esta investigación.

1. Introducción

La aparición del Internet en Colombia a mediados de los 90 puede ser catalogada como un boom tecnológico de grandes magnitudes. Dicha innovación tuvo un importante impacto que redefinió la manera en que las personas conciben su vida diaria. Debido al aumento de la accesibilidad y la evolución del contenido, el Internet tuvo efecto en todas las esferas del país, en los mercados nacionales e internacionales, en las empresas y en los hogares. De esta forma el Internet se convirtió en la puerta de acceso al conocimiento, la comunicación y el entretenimiento.

Estos cambios han sido de gran interés para los responsables de política pública, que ven en el Internet una herramienta para generar proyectos de impacto social y argumentan que la expansión de la cobertura de banda ancha es un mecanismo que lleva progreso a todas las regiones del país y en particular a las zonas periféricas. Es por eso que las autoridades gubernamentales han dedicado recursos importantes para proyectos que ayuden a acceder al Internet. De hecho, en el caso colombiano se invierte anualmente alrededor de un billón de pesos en proyectos que van desde llevar fibra óptica a toda la nación, hasta capacitar a las personas en el manejo de las tecnologías de la información y de la comunicación (MTIC 2011). En esta línea, la pregunta de investigación de este trabajo es analizar cuál es el impacto del Internet en el ingreso laboral para un país en desarrollo como Colombia.

En general, estudios empíricos y teóricos han encontrado que el mecanismo del impacto del Internet sobre el ingreso, se debe a que los trabajadores acumulan capital humano a través de esta herramienta, produciendo un aumento en su productividad y transmitida en los salarios. De hecho, no son necesarios grandes procesos de acumulación de capital humano, sino que el simple hecho de realizar un trabajo con mayor velocidad y eficiencia puede tener impacto sobre el ingreso. (Autor et al (2003), Bartel et al (2007) y Gutierrez (2011)). En cuanto la magnitud del coeficiente, se ha encontrado que el Internet tiene efectos positivos sobre diferentes indicadores a nivel país o individuo. Además se considera que las personas que no tienen acceso a Internet están en cierta desventaja de oportunidades comparadas con las que tienen acceso constante y de calidad en el servicio.

Ahora bien, el estudio del impacto del Internet en el ingreso tienen un problema de endogeneidad que lleva a resultados sesgados. Esto se debe a que los trabajadores pueden adoptar nuevas tecnologías porque son mejor pagos, en vez de recibir mejores ingresos por usar Internet (DiMaggio et al (2008)). Aunque la manera más apropiada de resolver este problema sería tener información tipo panel, tal que se pudiera observar cómo cambia el ingreso del individuo antes y después de usar Internet, esta información no es disponible para Colombia. Por lo tanto, siguiendo la revisión de literatura y el hecho de que usar Internet se debe a características observadas, se emplea el Propensity Score Matching (PSM) para eliminar el

problema de endogeneidad y encontrar estimadores consistentes, bajo el supuesto de que el balance en variables observadas implica un balance en variables no observadas. Este método ha sido utilizado en varios países del mundo, particularmente en donde se tiene información de encuestas de hogares (Navarro (2011), Benavente et al (2011) y DiMaggio et al (2008)). A partir de estos estudios, se plantea la primera hipótesis, que es encontrar el impacto del uso del Internet en el trabajo sobre el ingreso laboral en Colombia. Además esta primera conjetura plantea comparar los resultados con los otros países de la región y así mostrar que el efecto depende del capital humano y físico de cada país.

Después de estimar el modelo comparable con otros países, se procede a implementar mejoras en el modelo base de Navarro (2011), que incluye usar efectos fijos de municipio y de tiempo, otras variables de educación y socio-demográficas, así como una forma de emparejamiento más apropiada para la información disponible en Colombia. Con esta especificación se plantea la segunda hipótesis, que es la existencia de una prima mayor en el ingreso por manipular Internet en el trabajo, en vez de utilizarlo en el hogar. Aunque se debe tener en cuenta que el tipo de trabajador¹ juega un rol importante en los resultados, ya que el uso del Internet en el hogar es más importante para los empleados cuenta propia.

Con el modelo más completo según la información disponible se plantea la tercera hipótesis, que es encontrar un efecto positivo y significativo en las actividades más productivas realizadas en Internet, por ejemplo educación y aprendizaje. Sin embargo se puede considerar que no se halle ningún resultado porque el mercado no es capaz de diferenciar entre actividades productivas y ociosas. En esta hipótesis la frecuencia del uso de Internet tiene un papel fundamental, ya que es probable que para encontrar los resultados en el ingreso, se necesite varias horas de uso del Internet.

A partir de observar que usar Internet en actividades educativas debería tener un impacto sobre el ingreso, se plantea la cuarta hipótesis. Esta consiste en encontrar una prima mayor a medida que el Internet se complementa con mayores niveles educativos. Es decir, que las personas con más años de escolaridad manipulan el Internet con mayor productividad, tal que se incrementa el impacto en el ingreso laboral.

Como resultado de estas cuatro hipótesis, se logra mostrar si el lugar de acceso a Internet y/o las actividades que se realizan con esta herramienta son una señal más informativa para el mercado laboral. De manera que se debería priorizar el acceso en algunos lugares y fomentar actividades específicas que tengan impacto sobre la productividad y a través de este camino tener un efecto sobre el ingreso.

En este trabajo se utiliza la Gran Encuesta Integrada de Hogares (GEIH) entre 2009 - 2011, período en el cual se realizó el módulo sobre uso de las Tecnologías de la Información y

¹Tipo 1: Trabajadores Asalariados. Tipo 2: Trabajadores Cuenta Propia.

Comunicación (TIC). La aproximación principal es usar como variables explicativas las empleadas por la literatura, que consisten en las clásicas de la ecuación de Mincer (1995) y otras variables socio-económicas. Finalmente se implementa el PSM, usando como tratamiento el uso del Internet en el Trabajo² y como variable dependiente el logaritmo del ingreso laboral mensual.

La primera hipótesis se valida al estimar el modelo base de PSM, el cuál es el más comparable con los resultados de seis países en América Latina. Se encuentra que al igual que en los demás territorios el efecto difiere entre trabajadores asalariados y cuenta propia. Ordenando los resultados de mayor a menor, Colombia ocupa el puesto número 5 con un efecto de 0.21 para los asalariados y el puesto número 1 con un impacto de 0,43 para los cuenta propia³. Esto evidencia, que además de las diferencias en capital físico y humano de cada país, existe una gran heterogeneidad adentro de cada territorio que hace que el Internet sea más importante para unos trabajadores que para otros.

Luego de comparar los resultados de Colombia con los de países de la región, se utiliza la mejor especificación en cuanto a calidad de sus estimadores y del emparejamiento, para contrastar la segunda hipótesis, sobre la heterogeneidad del impacto según el lugar de acceso al Internet. Se encuentra que para los trabajadores asalariados el impacto de usar Internet en el trabajo (0.17) es mayor que el del uso en el hogar (0.10), situación similar para los cuenta propia (0.27 y 0.16 respectivamente). A partir de este mismo modelo, se responde la tercera hipótesis sobre las actividades que se realizan en Internet. Se distinguen dos tratamientos: primero el uso de Internet en educación y segundo el uso de esta herramienta para el entretenimiento. Aunque inicialmente no se encuentra resultados significativos para ninguna de las dos variables, la interacción de cada tratamiento con usar Internet diariamente muestra un efecto positivo y significativo para el uso de Internet en actividades educativas. La magnitud es de 0.04 para trabajadores asalariados y 0.07 para trabajadores cuenta propia, menor que la encontrada en cuanto al lugar de acceso. En cuanto a la cuarta hipótesis, se fija como grupo de control a las personas con secundaria incompleta y sin Internet, encontrando que a mayor educación y uso de Internet se incrementa la prima sobre los ingresos, tal que para los trabajadores asalariados con secundaria completa e Internet, el aumento del ingreso es de 0.12, mientras que con posgrado el efecto haciende a 0.57, así mismo el impacto se mantiene para los cuenta propia (0.17 y 0.47 respectivamente).

Lo que sigue en esta investigación es la sección 2 con la revisión de literatura, dividida entre estudios para América Latina y para el resto del mundo. La sección 3 expone la base de

²A lo largo del trabajo se cambia el tratamiento según las cuatro hipótesis.

³Los resultados de los 6 países son tomados de Navarro (2011), donde Brasil ocupa el primer puesto en magnitud en los trabajadores asalariados (0.30) y México el segundo puesto en cuanto a empleados cuenta propia (0.31).

datos, los argumentos para dividir la muestra en empleados asalariados y cuenta propia y las estadísticas descriptivas en cuanto a uso de Internet. La sección 4 desarrolla el argumento de por qué PSM es la mejor estrategia empírica debido a la información disponible para Colombia mientras que la sección 5 explica los resultados por tratamientos de lugar de uso y de manejo del Internet. Finalmente la sección 6 presenta las conclusiones y reflexiones.

2. Revisión de Literatura

El desarrollo de la tecnología ha traído grandes cambios en variables socio-económicas como el crecimiento del producto interno bruto (PIB), la tasa de desempleo, los salarios, la distribución del ingreso, la participación social y política. Por esta razón encontrar evidencia del efecto del Internet en estas variables es fundamental para realizar políticas públicas. De este modo, se ha abierto una agenda de investigación a nivel macro, donde el objeto de interés es el país y a nivel micro, donde se analizan los hogares y los individuos. En cuanto al análisis por país, Qiang et al (2009) encuentran que un aumento de 10 % de penetración de banda ancha incrementa entre 1.21 % - 1.38 % el crecimiento del PIB. Chu (2013) encontró que en términos per capita el aumento puede ser entre 0.57 - 0.63 pp. Freund et al (2001) muestran que el aumento de la penetración de banda ancha produce un incremento de 1.7 % de las exportaciones. Además Prasad et al (2001) muestran que el uso del Internet mejora la balanza comercial porque aumenta la integración y la productividad de las firmas.

En términos de análisis por individuo Acemoglu (1998) y Autor et al (1998) muestran que existe complementariedad entre trabajadores con educación alta y nuevas tecnologías, tal que el incremento en el uso del computador, aumenta la demanda por trabajadores asalariados y con esto su ingreso. En esta línea, Autor et al (2006) muestran que el mercado laboral de Estados Unidos se caracteriza por una "polarización" entre trabajos de salarios altos y de salarios bajos, a costa de una disminución de los trabajos de ingreso medio. Además los autores encuentran evidencia de que esto se debe a que los computadores son complementarios con tareas no cognitivas rutinarias y sustitutos con tareas rutinarias, por tanto los cambios en tecnología benefician a los trabajadores calificados y pueden perjudicar a los menos educados. En cuanto al uso del Internet, Goss et al (2002) encuentran una ganancia promedio de 13.5 % sobre el salario, para trabajos que usan esta herramienta, agregando que las industrias menos intensivas en uso de tecnología ofrecen una mayor prima sobre el ingreso para los trabajadores que hacen uso del Internet. En este mismo tema, DiMaggio et al (2008) muestra una ganancia de 20 % sobre el ingreso por hora, además el incremento en los salarios es gracias al Internet, no al simple uso de computadores. Forman et al (2012) analizan la inversión en Internet de las empresas de Estados Unidos en los noventa y encuentran una relación positiva y significativa con la tasa de empleo y el salario.

Ahora bien, en Latinoamérica existen varios estudios, por ejemplo Katz et al (2013) realizan una investigación en Ecuador y encuentran un efecto positivo de la penetración de Internet sobre el ingreso laboral de 3.67 % anual. Katz et al (2012) analizan el impacto económico de banda ancha en Panamá y muestran que la inversión en Internet contribuye en promedio a 0.44 % del PIB cada año. En general Benavides et al (2011) replica el modelo de Koutroumpis (2009) para 16 países de Latinoamérica y encuentran que un incremento de un punto porcentual de banda ancha, genera un aumento entre 0,03 % y 1,1 % del PIB. Además estos autores usan el modelo propuesto por Zahra et al (2008) para un panel de 18 países latinoamericanos y encuentran que si en 2009 el índice de infraestructura de telecomunicaciones ⁴ de Colombia (54.98) hubiese sido equivalente a la media para América Latina (60.1), se habría registrado un incremento entre 0,25 % y 0,46 % en la tasa de crecimiento del PIB per capita. Por otro lado, el proyecto líder del gobierno actual es Vive Digital que tuvo su primera etapa en 2011, la idea principal es mejorar la velocidad y la penetración del Internet, así como masificar este servicio. Katz et al (2011)⁵ encontraron que esta política pública tiene efectos positivos y significativos sobre variables como salarios y creación de nuevas empresas.

En cuanto al efecto del Internet en la productividad, Autor et al (2003) diseñan un modelo teórico que muestra que existe un sesgo de la tecnología hacia trabajadores calificados, de tal manera que una reducción en el precio del Internet,⁶ incrementa la demanda relativa por trabajadores con ventaja comparativa en tareas no rutinarias o calificadas, que se convierte en un aumento en sus salarios. Además Bartel et al (2007) analizan el mecanismo en el que las tecnologías de la información (TIC) promueve el crecimiento de la productividad de los trabajadores. Los autores muestran que la adopción de tecnologías en términos de capital coinciden con un incremento en las habilidades requeridas por los operadores de las maquinas, en temas técnicos y de solución de problemas, por lo que la inversión en tecnologías genera incentivos para que el capital humano mejore. En este mismo sentido Gutierrez (2011) hace un análisis con encuestas de firmas para Colombia y encuentran que las empresas con Internet tienen trabajadores más productivos, analizando esta variable como número de ventas y valor agregado por trabajador. Adicionalmente, la productividad laboral también puede ser originada en el hecho que las personas que gozan de acceso a Internet tienen mayores oportunidades. Por ejemplo Fountain (2003) argumenta que el uso del Internet es fundamental para acceder a servicios políticos y sociales que tienen un fuerte impacto en las habilidades de los ciudadanos, así como Tilley (2005) argumenta que las personas con difícil

⁴El índice de infraestructura de telecomunicaciones muestra la densidad de redes de comunicación por un millón de habitantes.

⁵ Katz et al (2011) usa información agregada a nivel departamental, controlando por el punto de partida en el nivel de desarrollo económico, el crecimiento de la población y el nivel de capital humano.

⁶Originalmente computadores en el lugar de trabajo.

situación social y económica, están inclusive peor si no tienen Internet.

Para estimar el efecto del Internet en el ingreso laboral se han usado modelos paramétricas como mínimos cuadrados ordinarios (MCO) y no paramétricas como Propensity Score Matching (PSM). En el primer caso, se ha hecho uso de la ecuación de Mincer (1958) que incluye la escolaridad o dummies relacionadas a esta variable y alguna medida de experiencia del individuo. Krueger (1993) usa MCO para estimar el impacto sobre el ingreso como función de la ecuación de Mincer, agregando una dummy igual a 1 si el individuo usa computador en el trabajo y luego realiza una variación incluyendo otra dummy cuando el individuo usa computador en el hogar⁷. En el segundo caso, Rosenbaum et al (1983) plantean el método de PSM para hacer inferencia causal. Este método ha sido utilizado para evaluar programas sociales y de política pública (Heckman et al (1998), Heckman et al (1999) y Dehejia et al (2002)). Particularmente, el PSM ha sido ampliamente empleado para encontrar el impacto del Internet en el ingreso laboral (DiMaggio et al 2008, Benavente et al (2011) y Navarro (2011)).

Particularmente, Navarro (2011) analiza el impacto del Internet sobre el salario para Brasil y Costa Rica en 2005, Chile en 2006, Honduras, México y Paraguay 2007, usando un modelo de emparejamiento por vecino más cercano y dividiendo la muestra entre trabajadores asalariados y cuenta propia. El autor encuentra que el uso del Internet tiene un efecto significativo en todos los países a excepción de Paraguay y el efecto es alrededor de 30 % para Brasil y Honduras, 25 % en Chile y Costa Rica y 17 % en México⁸. Estos resultados están por encima de la evidencia para Estados Unidos, donde estudios similares han encontrado que el efecto es alrededor de 15 % (Goss et al (2002), Freeman (2002) y DiMaggio et al (2008)). Para el caso de Colombia no se ha realizado un estudio usando encuestas de hogares que muestre el impacto del Internet sobre el ingreso laboral, por tanto esta investigación busca aportar novedosa evidencia para el país, además de hacer una comparación con vecinos de la región.

3. Datos y Estadísticas Descriptivas

En este estudio se utiliza la Gran Encuesta Integrada de Hogares (GEIH) para 2009 - 2011, la razón para elegir dicho período es que sólo en estos años se realizó el módulo de tecnologías de la información y de la comunicación. La base de datos consiste en una muestra mensual, por individuo y tiene representatividad a nivel nacional. Además esta encuesta tiene información del mercado laboral y de características socio-demográficas en Colombia.

En esta investigación únicamente se usan la población ocupada en tiempo completo y en

⁷Otros autores que han utilizado una especificación similar son Goss et al (2002) y Freeman (2002).

⁸Resultado de Navarro (2011) para empleados hombres y asalariados.

cascos urbanos⁹. Adicionalmente sólo se utiliza información para los trabajadores asalariados y cuenta propia. El primer grupo incluye empleados del gobierno, del sector privado, empleados domésticos y jornaleros¹⁰, mientras que en el segundo grupo están solamente los empleados cuenta propia, y se excluyen trabajadores familiares sin o con remuneración. Al mismo tiempo todas las estimaciones incluyen un factor de expansión mensual, el cual muestra la representación de cada observación de la muestra en la población total.

La variable de análisis es el logaritmo del ingreso laboral mensual, que es calculada sumando los ingresos por primera y segunda actividad económica. De acuerdo con la literatura se eliminan posibles observaciones atípicas de la variable dependiente, por tanto se excluye de la muestra el 1 % de los más ricos y el 1 % de los más pobres y se restringe la muestra entre 18 y 65 años (Navarro (2011) y DiMaggio et al (2008)).

El cuadro 1 muestra las estadísticas descriptivas de todas las variables usadas en esta investigación y son divididas en grupos con el objetivo de mostrar con mayor claridad, el proceso de selección del mejor modelo para realizar el Propensity Score Matching. En cuanto a lugar de uso, se consideran 4 tratamientos que son: usar Internet en el Trabajo, en el Hogar, en ambos lugares y en al menos uno de estos sitios. Por su parte, en la actividades realizadas con esta herramienta se utiliza como tratamiento; primero, el uso de Internet para educación y segundo, usar el servicio en actividades de entretenimiento¹¹. Se debe aclarar que las variables no son excluyentes, es decir, una persona puede acceder a Internet en más de un lugar y lo puede usar en más de una categoría. Además, se tiene información de cuantas veces se usa el Internet, diariamente, semanalmente, mensualmente y menor a una vez al mes. Estas variables serán fundamentales para analizar cómo la intensidad de uso juega un papel relevante en la acumulación de capital humano.

Por otro lado, el modelo es separado en dos muestras, una sólo para hombres y otra para ambos géneros. En general los estudios se han enfocado en el análisis para hombres¹², debido a que estimar algunas variables puede ser más complicado para las mujeres¹³, por lo que la idea es estimar los mismos modelos con las dos muestras y comparar los resultados, además la última especificación incluye una dummy para el género. A su vez, la base de datos es dividida entre trabajadores asalariados y trabajadores cuenta propia, ya que estos dos grupos

⁹Navarro (2011) y Benavente et al (2011), solo se enfocan en cabeceras por tanto para hacer el estudio comparable con otros países se debe seguir en esta línea.

¹⁰Empleados asalariados según definición del Departamento Administrativo Nacional de Estadística (DA-NE). Además de acuerdo con esta entidad los trabajadores familiares no son incluidos ni en empleados asalariados, ni en cuenta propia.

¹¹Textualmente se le pregunta al encuestado si usa Internet para juegos, bajar música entre otros.

¹²Ver Goss et al (2002), DiMaggio et al (2008), Benavente et al (2011) y Navarro (2011).

¹³Hotz et al (1988) muestran que el ciclo de la fertilidad puede tener impacto sobre la participación laboral de las mujeres y la experiencia laboral. Estas variables no son observada en la actual base de datos.

tienen diferentes características socio-demográficas y disparidad en cuanto al lugar de uso del Internet.

La figura 1 (Panel A) muestra el uso promedio del Internet en el trabajo entre los tipos de empleados. Se puede observar que los asalariados (0.59) usan más del doble el Internet en el trabajo que los cuenta propia (0.28). Además la figura 1 (Panel B) divide a los cuenta propia de acuerdo a si trabajan en el hogar o no. Se evidencia que para los cuenta propia que trabajan en el hogar el uso del Internet en el trabajo es menos del 10 %, por lo que el acceso al Internet se realiza principalmente desde la vivienda. De igual manera, el cuadro A1 del apéndice, muestra la diferencia de medias entre los dos tipos de trabajadores, sobre la variables más relevantes para el estudio, encontrando que todas las diferencias son estadísticamente significativa. Por este motivo la interpretación del impacto del Internet sobre el ingreso laboral depende del tipo del empleado, lo que justifica dividir la muestra. De hecho, la decisión de usar Internet es diferente según el tipo de trabajador, ya que los asalariados no escogen directamente tener Internet en el trabajo y más bien la decisión depende de la firma, mientras que los trabajadores cuenta propia si toman la decisión y tienen que pagar por este servicio. Considerando esta evidencia se analiza cada tipo de trabajador por separado y se analizan las cuatro hipótesis.

Ahora bien, el sector económico y la ocupación son variables fundamentales para este estudio ya que permiten comparar a las personas en la misma celda de oficio y actividad económica. La figura A1 muestra como los sectores de organizaciones gubernamentales (0.81), administración pública (0.79) e intermediación financiera (0.77) son los que más usan Internet en el trabajo. A su vez, la variable de ocupación contiene 99 categorías sin ningún tipo de orden ni categorización, por lo tanto de manera descriptiva se creó una variable llamada Trabajador Calificado, usando el Diccionario de Títulos Ocupacionales (BLS, (2010)). A través de dicho ejercicio, la figura A2 muestra como los trabajadores calificados (0.69) usan Internet mucho más intensivamente que los no calificados (0.29). En las estimaciones se incluyen los 17 sector económico y las 99 ocupaciones como efectos fijos, ya que esto hace que los resultados sean robustos a la definición de trabajador calificado.

4. Estrategia Empírica

La forma ideal de analizar el impacto del uso del Internet en el ingreso laboral sería con datos tipo panel, donde se tuviera al mismo individuo antes y después de acceder a este servicio. Una de las ventajas de esos datos es que permiten, luego de controlar por un conjunto de variables, observar si existe un cambio promedio en el ingreso. Sin embargo, este tipo de información no es disponible en Colombia. Incluso no se podría construir un panel

dinámico apropiado porque no es posible tener rezagos en el tiempo para el individuo i en variables endógenas (modelos autoregresivos) y/o exógenas (modelo de rezagos distribuidos), debido a que ningún individuo se repite en el tiempo¹⁴.

En cuanto al problema de sesgo de selección, que hace que los estimadores de mínimos cuadrados ordinarios (MCO) no sean insesgados, se podría considerar el método de variables instrumentales, que corrige el sesgo de selección basado en variables no observadas, usando instrumentos que predicen el tratamiento pero no el logaritmo del ingreso. Sin embargo según la revisión de literatura, tener Internet en el trabajo no depende de características no observables cuando se dispone de información de encuestas de hogares¹⁵. Además existe una dificultad en encontrar un buen instrumento que tenga variación a nivel de municipio y de tiempo.

Otro método para solucionar el sesgo de selección es el Propensity Score Matching (PSM), que es muy útil cuando la información no proviene de una asignación aleatoria del tratamiento (Internet en el trabajo), como es en este caso. Intuitivamente, el método de emparejamiento permite encontrar para cada individuo con Internet (tratado), otro “idéntico”, excepto por la participación en el programa (control). El supuesto fuerte de esta estrategia de estimación es que usar Internet en el trabajo no debe estar determinado por variables no observadas, sino que se atribuye exclusivamente a variables observadas del individuo (independencia condicional) (Bernal et al (2012)). Para lograr cumplir esta condición, la base de datos tiene una fuente amplia de información tanto económica como demográfica, por lo que se dispone de las características necesarias para conseguir estimadores consistentes. De hecho, el supuesto es que las brechas no observadas se cierran al mismo tiempo que las diferencias en variables observadas¹⁶. Adicionalmente, este método es no paramétrico, entonces no se deben hacer supuestos sobre la forma funcional, a diferencia de los modelos de regresión lineal¹⁷.

Ahora bien, para abordar las cuatro hipótesis de esta investigación, primero se estima una diferencia de medias entre los usuarios de este servicio en el trabajo y los que no lo usan, discriminando por niveles educativos. Aunque la magnitud de la diferencias disminuye cuando aumenta el nivel de la educación, dicha diferencia siempre es estadísticamente significativa y

¹⁴De acuerdo con Arellano et al (1991) es necesario tener variables rezagadas de X y/o Y para construir un panel dinámico. Sería posible si se considera a i como el municipio y hacer regresiones a nivel macro, pero con la desventaja de perder variabilidad de la información, que es lo valioso de trabajar con encuestas de hogares.

¹⁵Benavente et al (2011), Navarro (2011) y DiMaggio et al (2008) usan encuestas de hogares para diferentes países del mundo. Los dos primeros usan datos de corte transversal.

¹⁶Dehejia et al (2002) muestran que la dimensionalidad entre características observadas es alta, es decir que con una variable binaria se pueden cerrar las brechas entre grupos de otra variable.

¹⁷Los modelos de regresión lineal incluye MCO y variables instrumentales.

a favor de los usuarios del Internet, por ejemplo, personas con secundaria incompleta tienen una prima extra sobre el ingreso laboral de 0.53 puntos porcentuales, mientras que con posgrado la prima es de 0.34 (Figura 2). A pesar de esta clara evidencia, el test de medias no controla por características socio-demográficas, por esta razón el segundo paso para abordar la pregunta de investigación es estimar la ecuación de Mincer (1958), que incluye una dummy igual a uno si el individuo usa Internet en el trabajo, de tal forma que la función para estimar se puede expresar como:

$$\ln Y_{it} = \beta_0 + \beta_1 Educacion + \beta_2 Edad + \beta_3 Edad^2 + \theta Internet en el Trabajo + X_{it} + \gamma_t + \delta_j + \eta_{it} \quad (1)$$

Donde la variable dependiente es el logaritmo del ingreso laboral en el individuo i en el tiempo t , *Educacion* es una dummy igual a 1 para doce y más años de educación, *Edad* muestra la edad de la persona, X_{it} es un conjunto de variables socio-demográficas y de uso de tecnología¹⁸ y finalmente γ_t y δ_j son efectos fijos de tiempo y municipio respectivamente. Estos dos últimos componentes son esenciales, ya que el primero captura cualquier heterogeneidad observada o no observada a nivel municipal si esta no cambia en el tiempo y el segundo captura choques temporales que sean agregados en el tiempo.

Debido al problema de endogeneidad la ecuación 1 sobrestima el impacto del Internet en el trabajo porque en ese modelo no es claro si los trabajadores tienen mayor ingreso por usar Internet o si la causalidad es inversa. Por tanto, el tercer paso, que es la estrategia de identificación, consiste en usar el método de PSM, planteado por primera vez por Rousenbaum et al (1983) en estudios de efectos causales. El estimador de emparejamiento puede ser escrito como:

$$E[Y_{1i} - Y_{0i} | D_i = 1] = \sum_x \theta_x P[(X_i = x) | D_i = 1] \quad (2)$$

Donde Y_{1i} es el ingreso de los individuos i tratados, Y_{0i} es el ingreso de los no tratados, D_i es una dummy igual a 1 cuando el individuo usa Internet en el trabajo, X_i es un conjunto de características socio-económicas del individuo i y $P[(X_i = x) | D_i = 1]$ es la probabilidad estimada de recibir el tratamiento dadas las características observables X_i (Propensity Score). Es necesario resaltar que el modelo de emparejamiento encuentra el efecto promedio del tratamiento sobre los tratados (θ_{ATT}). Entonces, este parámetro muestra cuánto gana o pierde un trabajador como consecuencia de usar Internet, mientras que el efecto promedio del tratamiento (θ_{ATE}) muestra cuánto gana o pierde, cualquier persona, no sólo trabajadores, como consecuencia de usar el Internet (Angrist et al (2008)). La forma de implementar este método es:

¹⁸Incluye efectos fijos por las 99 ocupaciones en las que son divididos los trabajadores y los 17 sectores de la economía de acuerdo con la clasificación industrial internacional uniforme de todas las actividades económicas (CIIU). Ambas divisiones son realizadas por el DANE.

1. Se construye el Propensity Score (PS), estimando un modelo logit de la probabilidad de ser tratado, es decir usar Internet en el trabajo, como función de las características individuales, como está escrito en la ecuación 3.

$$P(\text{Internet en el Trabajo})_{it} = \beta_0 + \beta_1 \text{Educacion} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 + X_{it} + \gamma_t + \delta_j + \eta_{it} \quad (3)$$

2. Se emparejan el PS entre individuos tratados y de control y se estima el efecto del Internet en el trabajo sobre el ingreso laboral, siguiendo la ecuación 4.

$$\text{Ln}Y_{it} = \beta_0 + \beta_1 \text{Educacion} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 + \theta_x \text{Internet en el Trabajo} + X_{it} + \gamma_t + \delta_j + \eta_{it} \quad (4)$$

El hecho de que el PS incluya efectos fijos de municipio y de tiempo hace que se reduzca el sesgo por variable omitida. El primer efecto fijo, entre otras cosas, controla que la probabilidad se vea afectada por variables como ubicación, clima, altura entre otras, que pueden ser muy relevante a la hora de determinar el tratamiento (usar Internet en el trabajo), ya que las ciudades principales tienen mayor cobertura y calidad del servicio. Por otro lado incluir en X_{it} sector económico y ocupación laboral hace que se comparen personas en trabajos y en jerarquías laborales similares¹⁹.

El objetivo inicial es responder la primera hipótesis de cuál es el impacto del uso del Internet en el ingreso laboral, para el caso colombiano, comparando con otros países de América Latina. Por esta razón se sigue como línea base el estudio de Navarro (2011) para seis países en esta región y dado que los resultados de los otros países son generados a partir de datos de corte transversal para un solo año, se trabaja solamente con el año 2009 en Colombia²⁰. Además se estima el modelo de PSM usando el emparejamiento de vecino más cercano, el cual empareja cada individuo del grupo del tratamiento (PS_i) con un individuo en el grupo de control (PS_c), de acuerdo al PS y siguiendo una distancia euclidiana que puede ser escrita como:

$$d_e(i, j) = \sqrt{\sum_{h=1}^p (PS_i - PS_c)^2} \quad (5)$$

Luego de emparejar a los individuo se estima el θ_{ATT} , que se expresa de la siguiente manera:

$$\theta_{ATT} = \sum_{i=1}^I \text{mean}[(Y_i|D_i = 1) - (Y_{c(i)}|D_i = 0)] \quad (6)$$

¹⁹Por ejemplo se emparejan arquitectos e ingenieros (ocupación) en la rama de la educación (sector económico) que usen y no el Internet en el trabajo, además de otras variables características. Este mismo ejercicio se repite para las demás ocupaciones y sectores económicos.

²⁰Este año es el más cercano a la información de los otros países.

Donde I es el número de individuos en el grupo de control, D_i es igual a 1 si el individuo i es tratado, C_i es el conjunto de vecinos más cercanos en el grupo de control al individuo i , por lo que el coeficiente θ_{ATT} es un promedio entre los resultados de ambos grupos (Bernal et al (2012)).

Ahora bien, dada la información disponible para Colombia la especificación se puede mejorar para explicar más apropiadamente las hipótesis de lugar y manejo del Internet. Para dicho proceso, lo primero que se realiza es usar el periodo de información completo, tal que se trabaja con observaciones desde 2009 a 2011. Segundo, se supera una de las principales críticas al trabajo de Navarro (2011), incluyendo efectos fijos de tiempo y municipio en todas las estimaciones. Tercero, debido a que el estudio base incluye una dummy que toma el valor de 1 para doce o más años de escolaridad, pero en Colombia el aprendizaje del Internet es en secundaria, se hace necesario incluir un conjunto de variables más amplio que analice mejor la diferencia de completar este grado. Una alternativa sería incluir 8 dummies²¹ por nivel educativo completo e incompleto, sin embargo cuando hay muchas variables es complicado emparejar a los individuos porque se corre el riesgo de que en alguna celda no se encuentre individuos de tratamiento y de control (Dehejia et al (2002)). En este sentido, se consideran 5 dummies, tal que una agrupe a los individuos con educación por debajo de secundaria completa²². Finalmente, el cuarto aporte al modelo base, es incluir otras variables relevantes en la probabilidad de estar en el grupo de tratamiento, tales como estado civil, afiliación a régimen subsidiado de salud y cotización a pensión. Por tanto el modelo final consta de un nuevo Propensity Score que se puede escribir como:

$$P(Interne nel Trabajo)_{it} = \beta_0 + \beta_1 Conjun to Educacion + \beta_2 Edad + \beta_3 Edad^2 + \hat{X}_{it} + \gamma_t + \delta_j + \eta_{it} \quad (7)$$

Donde $ConEducacion$ es un conjunto de dummies de niveles educativos alcanzados, $\hat{X}_{it}$ son las características socio demográficas ya mencionadas, más las otras variables relevantes. Luego de estimar el PS, se calcula el efecto del Internet en el Trabajo en el ingreso laboral, con la siguiente ecuación:

$$LnY_{it} = \beta_0 + \beta_1 Conjun to Educacion + \beta_2 Edad + \beta_3 Edad^2 + \theta_x Interne nel Trabajo + \hat{X}_{it} + \gamma_t + \delta_j + \eta_{it} \quad (8)$$

Por otro lado, para que el método de emparejamiento funcione apropiadamente y se puedan encontrar estimadores insesgados, las brechas en características entre el grupo de tratamiento

²¹Estas 8 dummies son: Ninguna educación, primaria incompleta, primaria completa, secundaria incompleta, secundaria completa, universidad incompleta, universidad completa y posgrado.

²²Esta agrupación mantiene la utilidad de analizar la secundaria completa e incompleta y se basa en DiMaggio et al (2008).

y de control no deben ser estadísticamente significativas. Aunque Navarro (2011) usa el algoritmo del vecino más cercano, debido a la mejora del modelo base y en especial a la valiosa inclusión de efectos fijos, es necesario utilizar la distancia de Mahalanobis. Esta medida es mucho más flexible que el algoritmo del vecino más cercano, por lo que es más fácil emparejar individuos del grupo de control con los de tratamiento. La distancia de Mahalanobis puede ser escrita como:

$$d_e(i, j) = \sqrt{\sum_{h=1}^p (PS_i - PS_c) S(X_{ic})^{-1} (PS_i - PS_c)'} \quad (9)$$

La ventaja de este algoritmo de emparejamiento es que tiene en cuenta la matriz de varianza y covarianza entre las variables explicativas ($S(X_{ic})^{-1}$), tal que resulta ser más informativa en cuanto a emparejar a los individuos más parecidos²³.

Para responder la segunda hipótesis sobre la variación de la magnitud del efecto dependiendo el lugar donde se usa el Internet, se estima el modelo de PSM, con el algoritmo de Mahalanobis, cambiando el tratamiento por usar Internet en el hogar, en el trabajo y en el hogar y al menos en alguno de estos sitios. Por tanto se estima la siguiente ecuación:

$$\begin{aligned} \ln Y_{it} = & \beta_0 + \beta_1 \text{ConjuntoEducacion} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 \\ & + \theta_x \text{LugardeUso} + \hat{X}_{it} + \gamma_t + \delta_j + \eta_{it} \end{aligned} \quad (10)$$

En cuanto a la tercera hipótesis sobre el manejo del Internet y el impacto en el ingreso, se utiliza el modelo final, seleccionando como tratamiento usar Internet en diferentes actividades, es decir estimando la siguiente ecuación:

$$\begin{aligned} \ln Y_{it} = & \beta_0 + \beta_1 \text{ConjuntoEducacion} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 \\ & + \theta_x \text{ManejodeInternet} + \hat{X}_{it} + \gamma_t + \delta_j + \eta_{it} \end{aligned} \quad (11)$$

Además, es importante la frecuencia en la que se usa el Internet, ya que la intensidad con la que se maneja este servicio puede ser clave para que el mercado laboral logre identificar a las personas que usan Internet de la forma más apropiada, por lo que se utiliza una interacción entre usar Internet para educación con una dummy igual a 1 cuando el individuo usa esta

²³El método del vecino más cercano forma una círculo sobre el soporte común y empareja a los tratados y no tratados. El método de Mahalanobis forma un ovalo sobre este soporte, tal que usa un número mayor individuos control, para emparejarlos con tratados.

herramienta diariamente. Por tanto se estiman la ecuación:

$$\begin{aligned} \ln Y_{it} = & \beta_0 + \beta_1 \text{ConjuntoEducacion} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 \\ & + \theta_x \text{ManejodeInternet} * \text{UsoDiario} + \hat{X}_{it} + \gamma_t + \delta_j + \eta_{it} \end{aligned} \quad (12)$$

Esta investigación no quiere presentar al Internet como una escape a la educación, en el sentido que erróneamente se podría considerar que no es necesario asistir a un centro educativo si se maneja Internet en el trabajo. Tampoco pretende negar la extensa literatura que relaciona la educación y el ingreso²⁴. Por esta razón se plantea la hipótesis 4, con un modelo donde el tratamiento es la interacción entre usar Internet en el trabajo y una dummy de nivel educativo, de tal manera que se pueda observar que a medida que aumenta la escolaridad el efecto de esta herramienta es mayor. Por lo que se presenta evidencia de que el Internet, potencializa las habilidades que se aprenden en los centros educativos. La ecuación que se estima es:

$$\begin{aligned} \ln Y_{it} = & \beta_0 + \beta_1 \text{ConjuntoEducacion} + \beta_2 \text{Edad} + \beta_3 \text{Edad}^2 \\ & + \theta_x \text{InternetenelTrabajo} * \text{NivelEducativo} + \hat{X}_{it} + \gamma_t + \delta_j + \eta_{it} \end{aligned} \quad (13)$$

Para estimar esta ecuación se fija el grupo de control como los individuos sin Internet y con secundaria incompleta.

5. Resultados

Debido a que no es clara la causalidad entre el uso del Internet y el ingreso, mínimos cuadrados ordinarios (MCO) estima coeficientes que sobrestiman el efecto. Por esta razón, se plantea el Propensity Score Matching (PSM) para lograr eliminar el problema de endogenidad y encontrar estimadores no sesgados. Además se mostró que este modelo es la mejor estrategia empírica debido a la información disponible para Colombia.

Ahora bien, para garantizar que PSM es un método valido para el análisis de las cuatro hipótesis planteadas, se debe asegurar que existe un soporte común entre los individuos tratados y de control. Este soporte se define como la región en la que el PS_i y el PS_c se intersectan. El tamaño del soporte es importante porque en esta región se emparejan a los individuos de control con los de tratamiento. La figura 3 muestra el Propensity Score (PS)

²⁴Ver Mincer (1958), Morgan et al (1963), Link et al (1975), Blundell et al (2005) y Dearden et al (2009).

para individuos tratados (i) y de control (c), estimado con la ecuación 7, para trabajadores asalariados. Se puede observar que el soporte común incluye al 90 % de la muestra y las observaciones que no se incluyen pertenecen al grupo de tratamiento. La figura A3 muestra que existe un soporte común más amplio para los empleados cuenta propia, en comparación con los asalariados, lo que valida que el PSM es un método útil para ambos tipo de trabajadores y para abordar las cuatro hipótesis.

La figura 4 muestra la distribución de la escolaridad y la edad antes y después de usar el emparejamiento de la distancia de Mahalanobis. Se evidencia que antes de emparejar a los individuos las distribuciones son muy diferentes entre el grupo de tratamiento y de control. Sin embargo después de usar el PSM, las distribuciones se sobreponen, es decir que las brechas en estas características se cierran. Por otro lado, aunque la variable experiencia²⁵ no se incluyó en el PSM, se puede observar como también las diferencias se cierran entre tratados y controles. Lo anterior es muestra de que al balancear variables observadas, también se logra cerrar la brecha en otras características, incluso si estas no se usan directamente en la especificación. La figura A4 muestra resultados análogos para empleados cuenta propia, donde se sigue manteniendo la efectividad de la distancia de Mahalanobis para cerrar las brechas entre las variables.

En cuanto al tipo del emparejamiento, el cuadro 2 muestra la diferencia de medias entre tratados y no tratados, antes y después del emparejamiento, para empleados asalariados. Se puede observar que antes de usar el método del PSM, las diferencias entre el grupo de tratamiento y el de control son estadísticamente significativas. Luego de utilizar el método del vecino más cercano, 8 de las 10 diferencias de medias son diferentes a cero. Mientras que, al usar distancia de Mahalanobis, se halla que las 10 diferencias de medias se cierran, por lo que este método logra corregir el problema de endogeneidad. Por lo tanto, se puede observar que el emparejamiento con la distancia de Mahalanobis usando efectos fijos es la mejor especificación en cuanto a emparejar el grupo de tratados y de control. El cuadro A2 en el apéndice, muestra resultados análogos para los empleados cuenta propia.

5.1. Colombia en América Latina

El cuadro 3 muestra el coeficiente $\hat{\theta}$ del efecto del Internet en el trabajo sobre el ingreso laboral estimado por MCO. Se aprecia que el impacto es mucho más grande para los empleados cuenta propia (0.47) que para los asalariados (0.23). Además separar la muestra entre hombres y toda la población evidencia un interesante resultado y es que al incluir a las mujeres, el coeficiente cae para ambos trabajadores (0.43 y 0.21 respectivamente).

²⁵Siguiendo a Benavente et al (2011), la variable experiencia se genera como la edad menos los años de escolaridad menos 6, este último componente se debe a que la persona empieza a estudiar a los 6 años

Además el cuadro 3 muestra los resultados del modelo base, que se construye a partir de usar las variables y el algoritmo de emparejamiento de Navarro (2011), solamente para el año 2009. De esta manera se responde la primera hipótesis, ya que se estiman coeficientes que pueden ser comparados con los resultados de otros países de América Latina y se valida la idea de que el uso del Internet en el trabajo tiene un impacto positivo y significativo sobre el ingreso laboral. Por lo tanto el estimador para los trabajadores asalariados en Colombia (0.23) está entre México (0.17) y Chile (0.26) y lejos de los de mayor impacto, que son Honduras (0.30) y Brasil (0.29). Sin embargo en cuanto a los empleados cuenta propia, Colombia (0.47) muestra la mayor magnitud seguido por México (0.32), Honduras (0.31), Brasil (0.21) y Chile (0.18)²⁶.

5.2. Lugar de Uso

El cuadro 4 muestra el θ_{ATT} para el modelo final, que incluye efectos fijos de tiempo y municipio, otras variables de escolaridad y socio-demográficas, así como el emparejamiento por la distancia de Mahalanobis. Se encuentran resultados consistentes que validan la hipótesis de que el Internet en el Trabajo tiene un impacto significativo y positivo en el ingreso laboral. La magnitud del coeficiente es de 0.17 para trabajadores asalariados y 0.27 para cuenta propia. En cuanto a la muestra con ambos géneros se puede observar que el coeficiente es 0.12 y 0.32 respectivamente. Es decir, que se reduce la magnitud del coeficiente con respecto al modelo base porque en la especificación final se disminuye el problema de tener sesgo por variable omitida.

Ahora bien, para calcular que tan robusto es el resultado, Sianesi et al (2003) propone eliminar algún porcentaje del soporte común y analizar como cambian los resultados. El cuadro A3 muestra que los resultados se mantienen más o menos constantes al eliminar el 5 %, 10 % y 20 % del soporte común, lo que muestra que el impacto del Internet en el trabajo sobre el ingreso laboral es robusto a variación en el PS.

Los efectos encontrados hasta ahora son desarrollados en esta investigación tomando como tratamiento usar Internet en el Trabajo. La segunda hipótesis plantea que existe un efecto heterogéneo que depende de dónde se accede al servicio. En ese sentido, se procede a tomar como tratamiento usar Internet en el Hogar, en ambos lugares y al menos en alguno de estos sitios. Los resultados de este nuevo ejercicio se muestran en el cuadro 4 estimando con PSM el modelo final (ecuación 10). Se observa que para ambos tipos de trabajadores el impacto del uso de Internet en el hogar es mucho menor en magnitud comparado con

²⁶Los cálculos de cualquier país latinoamericano, distinto a Colombia son tomados de Navarro (2011). Paraguay y Costa Rica no son mencionados porque no encuentran efectos significativos en los dos grupos de trabajadores. Además este modelo solo incluye en la muestra hombres.

el uso en el trabajo. A pesar de esto, el Internet en el hogar tiene una mayor importancia para los trabajadores cuenta propia (0.16), comparados con los asalariados (0.10) y el efecto se mantiene al incluir ambos géneros. Además, se puede observar complementariedad entre lugares de uso para los cuenta propia, ya que usar Internet en los dos lugares o al menos en uno aumenta el efecto a 0.30. En el caso de los asalariados, se podría pensar en un efecto inverso o en la inexistencia del efecto, ya que en estos trabajadores cae el impacto a 0.15 si usa Internet en ambos lugares y se mantiene si se utiliza en al menos un lugar. El cuadro A4 muestra para estos dos últimos coeficiente dos pruebas: primero un test de medias, que compara si la diferencia de los coeficientes en promedio es igual a cero y segundo un test sobre las especificaciones, que analiza si ambos modelos son iguales o no, es decir que esta es una prueba F sobre todas las variables del modelo. Ambas pruebas no son estadísticamente significativas al 1 %, por lo que no es preciso decir que la diferencia es realmente importante.

5.3. Manejo del Internet

La tercera hipótesis plantea el hecho que las actividades que se realizan en el Internet tenga efecto en el ingreso, es decir que el mercado laboral es capaz de identificar o de diferenciar a los individuos que lleven a cabo tareas más productivas que otras. Si bien esta pregunta puede ser muy ambiciosa dada la disponibilidad de la información, al menos se encuentra un indicio de lo que se podría hallar con una encuesta más detallada sobre el uso del tiempo de los individuos. El cuadro 5 muestra los resultados para usar el Internet en entretenimiento o en educación. Si bien no se encuentra ningún coeficiente significativo, este panorama cambia al incluir la interacción con la variable de frecuencia de uso del Internet, que además de arrojar coeficientes positivos y significativos, también muestra un impacto del uso de Internet para actividades de educación. De nuevo se puede observar que el efecto es mayor para empleados cuenta propia (0.07) que para los asalariados (0.04). En cuanto a la muestra total se observa que la magnitud se mantiene para los asalariados y cae marginalmente para los cuenta propia (0.06). Además la tabla A5 muestra como las personas que manejan esta herramienta semanalmente también tienen un impacto positivo y significativo sobre el ingreso laboral, sin embargo la magnitud es coeficiente es menor que el reportado con uso diario en esta tarea (0.02 para asalariados y 0.03 para cuenta propia).

Este hallazgo es valioso porque muestra que si bien el manejo del Internet no tiene efecto sobre el salario, la intensidad con la que se realicen las actividades es fundamental para lograr que un aumento en capital, que se pueda reflejar en un mayor ingreso laboral.

5.4. Educación e Internet

El último resultado de manejo de Internet sugiere que la educación es clave para determinar el efecto del Internet en el ingreso laboral. El cuadro 6 muestra los resultados de las estimaciones cuando el tratamiento es la interacción entre usar Internet en el trabajo y algún nivel de educación. El grupo de control en todas las estimaciones es las personas con secundaria incompleta y sin Internet. Se observa que la hipótesis 4 se valida, debido a que el efecto de este servicio aumenta con el nivel de escolaridad, tal que para los trabajadores asalariados con secundaria completa el efecto es 0,12 y con posgrado el efecto es 0,57, mientras que para los cuenta propia el efecto es 0.17 y 0.47 respectivamente. La muestra completa revela resultados iguales en dirección y similares en magnitud.

Un resultado muy interesante del cuadro 6 es que presenta el menor coeficiente para los asalariados que tienen menor educación. Esto se debe a que este tipo de trabajadores sufren de un castigo económico muy grande, que ni siquiera el Internet logra solucionar. Por el contrario los trabajadores asalariados más educados reciben una prima mayor cuando usan Internet, ya que gozan el coeficiente más alto de todo el estudio. En el caso de los empleados cuenta propia, la penalidad por la menor educación no es tan importante como para los asalariados, sin embargo se mantiene que el Internet agrega una prima en el ingreso laboral, cuando las personas avanzan en el nivel educativo.

5.5. Eliminación de Sesgo

Finalmente para mostrar que MCO efectivamente no controla el problema de endogeneidad y estima coeficientes sesgados, se muestran cada estimación con este método y con PSM. Al estimar la hipótesis 2, sobre efectos heterogéneos dependiendo del lugar de uso, se puede observar que en general los coeficientes de MCO son casi siempre mayores a los de PSM, a lo sumo iguales (ver cuadro A6). Esta misma situación se presenta en la hipótesis 3, sobre manejo del Internet (ver cuadro A7) y en la hipótesis 4, acerca de la complementariedad que existe entre educación e Internet (ver cuadro A8). Por lo que efectivamente MCO no logra controlar el problema de endogeneidad entre usar Internet y tener mayor ingreso laboral, mientras que el modelo de PSM si corrige este efecto. Además, como se evidenció es la mejor estrategia empírica que se puede utilizar dada la información disponible en Colombia.

6. Conclusiones

Este estudio utiliza el modelo de PSM para encontrar estimadores insesgados. Además, esta es la mejor estrategia empírica que se puede emplear debido a que no se tiene información tipo panel para Colombia y que el sesgo de selección se debe a características que se pueden observar en la base de datos. A su vez, este modelo tiene la ventaja de balancear la muestra, incluso en variables que no se usan directamente en el emparejamiento, como se mostró con el caso de la experiencia. Se parte de un modelo base planteado por la revisión de literatura y se realizan modificaciones fundamentales para encontrar estimadores no sesgados y acorde a la realidad de este país. Estas mejoras son principalmente incluir efectos fijos, utilizar un conjunto de dummies de educación y usar el método de emparejamiento de Mahalanobis.

Por otro lado, se validan cuatro hipótesis: La primera, el Internet en el trabajo tiene un efecto positivo y significativo en Colombia y se comparan los resultados con otros países de la región. La segunda, que el efecto es heterogéneo dependiendo del lugar de acceso al Internet, siendo la magnitud más importante en el trabajo, que en el hogar. La tercera, las actividades realizadas con esta herramienta sólo tienen impacto sobre el Internet cuando se usan con frecuencia diaria. Finalmente, la cuarta es el Internet tiene un mayor impacto sobre el ingreso, a medida que aumenta el nivel educativo, siendo esto la conexión entre el mecanismo del aumento de capital humano y uso de esta herramienta.

En esencia la presente investigación más allá de encontrar evidencia única en esta clase de estudios para Colombia, muestra resultados valiosos para autoridades públicas y privadas. En el primer caso, se muestra que usar Internet en el trabajo tiene un mayor efecto que en el hogar, por lo que si es necesario priorizar el presupuesto nacional se debería empezar por asegurar el acceso para las firmas, sin olvidar que el hogar también presentó efectos positivos en el ingreso. En el segundo caso, este estudio muestra que las empresas deberían realizar inversiones en Internet y en otras tecnologías que hagan que los empleados sean más productivos y con esto aumenten los beneficios. Además, se encuentra que usar Internet en educación debería ser incentivado por los gobiernos locales o por las empresas privadas, ya que muestra efectos importantes sobre el ingreso. Finalmente se llega a la conclusión que educarse y usar Internet en el trabajo es la mejor forma de acceder a todos los beneficios que hacen que la productividad aumente y con esto el ingreso laboral.

7. Cuadros y Figuras

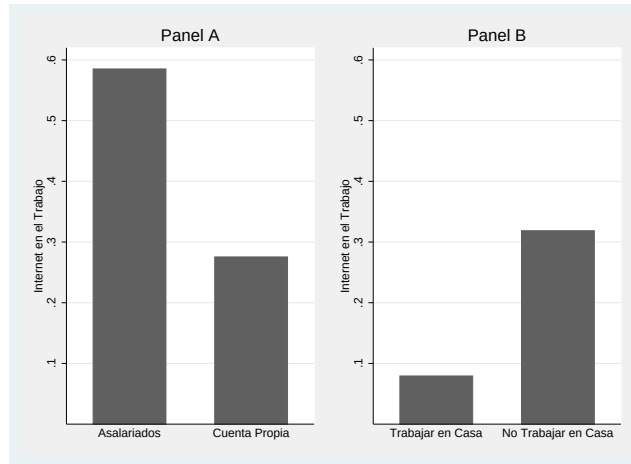
Cuadro 1: Estadísticas Descriptivas

Variable	Promedio	Dev. Est.	Min.	Max.
Variable Dependiente				
Logaritmo del Ingreso	13.53	0.90	0	15.61
Tratamientos				
Lugar de Uso				
Internet en el Trabajo	0.47	0.50	0	1
Internet en el Hogar	0.52	0.50	0	1
Internet en ambas partes	0.26	0.44	0	1
Internet en al menos un lugar	0.73	0.44	0	1
Manejo				
Educación	0.49	0.50	0	1
Entretención	0.55	0.50	0	1
Lucas Navarro (2011) - X_{it}				
Edad	32.91	9.74	18	65
Edad al cuadrado	1177.75	695.42	324	4225
12 o mas años de educación	0.65	0.48	0	1
Servicio telefónico	0.61	0.49	0	1
Tv por cable	0.99	0.12	0	1
Computador en el Hogar	0.61	0.49	0	1
Vivir en una Casa	0.54	0.50	0	1
Vivir en un Apt.	0.45	0.50	0	1
Casa propia	0.42	0.49	0	1
Casa en pago	0.40	0.49	0	1
Casa Arrendada	0.08	0.27	0	1
Trabajador Calificado	0.47	0.50	0	1
Actividad Económica	9.52	3.55	1	18
Ocupación	40.03	25.56	0	99
Variables Adicionales - $\hat{X}_{it}$				
Secundaria incompleta	0.12	0.13	0	1
Secundaria completa	0.34	0.47	0	1
Universidad incompleta	0.31	0.46	0	1
Universidad completa	0.22	0.41	0	1
Posgrados	0.12	0.32	0	1
Casado/a	0.47	0.50	0	1
Pago de Pensión	0.59	0.49	0	1
Salud Subsidiada	0.72	0.45	0	1
División de la Muestra				
Género	0.49	0.50	0	1
Trabajador Asalariado	0.59	0.49	0	1
Frecuencia de uso de Internet				
Diario	0.32	0.47	0	1
Semanal, pero no diario	0.23	0.42	0	1
Mensual, pero no semanalmente	0.05	0.21	0	1
Menos de mensual	0.01	0.10	0	1
Otras Variables				
Cuello blanco	0.48	0.50	0	1
Escolaridad	13.67	2.60	2	26
Experiencia	13.25	9.17	0	40

Número de observaciones 300,742.

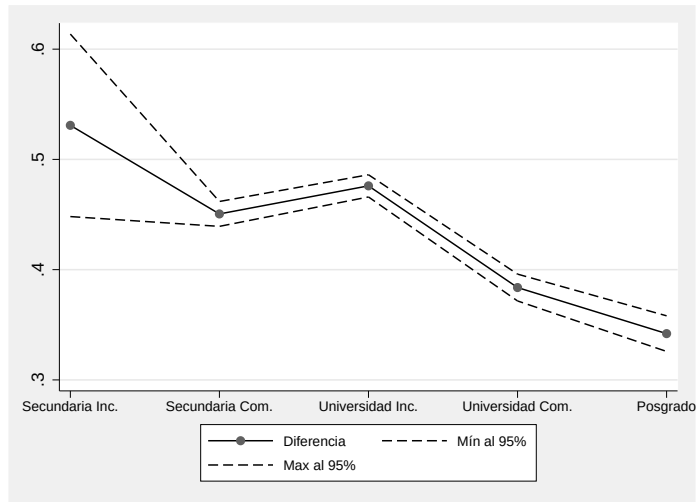
Las otras variables no se incluyen en las regresiones pero se utilizan de manera descriptiva en las secciones 3 y 5.

Figura 1: Internet en el Trabajo



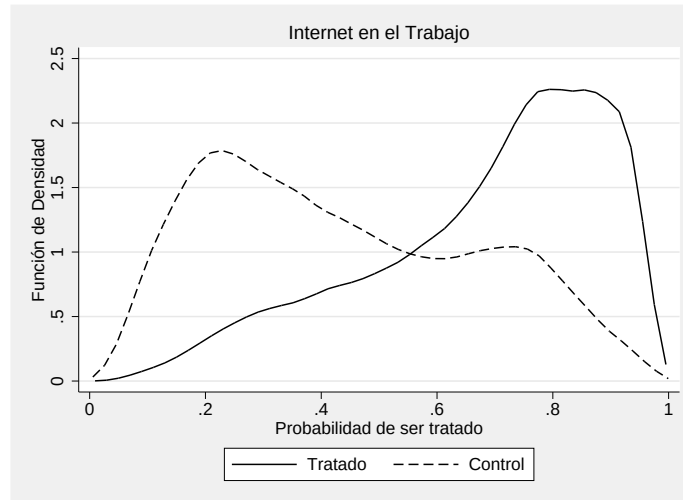
Panel A es usando toda la muestra y panel B es usando la muestra solo para cuenta propia

Figura 2: Diferencias en medias entre usuarios del Internet y no usuarios



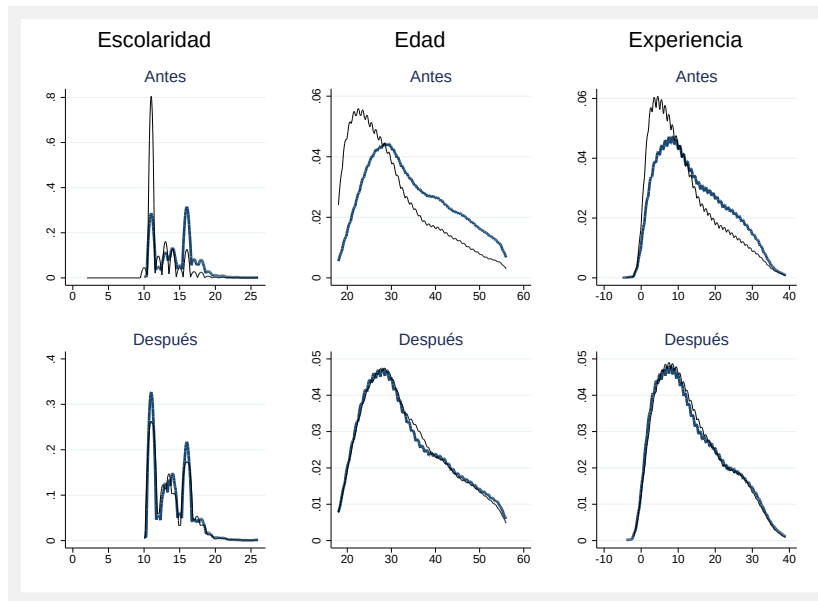
Min y max muestran respectivamente los intervalos de confianza al 95 %

Figura 3: Propensity Score para Empleados Asalariados



PS con el modelo final (Ecuación 7). El soporte común incluye el 90 % de los datos de la muestra. Las observaciones que no se incluyen pertenecen a el grupo de tratamiento.

Figura 4: Distribución antes y después del emparejamiento para Empleados Asalariados. Usando la distancia de Mahalanobis



Estimaciones usando el PSM (Ecuación 7 y 8). Línea Oscura grupo de Tratamiento. Línea clara grupo de control.

Cuadro 2: Diferencia de medias entre tratados y no tratados antes y después del emparejamiento. Modelo para empleados Asalariados

Variable	Antes (U)	Promedio		Diferencia	%Reducción del Sesgo
	Después (M)	Tratados	No Tratados		
Doce o más años de educación	U	0.74	0.41	0.33***	
	M-VC	0.74	0.74	0.00	99.40
	M-Maha	0.72	0.72	0.00	100.00
	M-Maha-EF	0.72	0.72	0.00	100.00
Edad	U	35.05	30.53	4.51***	
	M-VC	37.14	36.46	0.69***	84.80
	M-Maha	35.13	35.12	0.01	99.70
	M-Maha-EF	34.49	34.48	0.01	99.70
Edad al cuadrado	U	1316	1012	304.50***	
	M-VC	1472	1421	51.40***	83.10
	M-Maha	1322	1321	1.40	99.60
	M-Maha-EF	1277	1276	1.40	99.50
Teléfono Fijo	U	0.63	0.56	0.07***	
	M-VC	0.61	0.62	-0.01	90.70
	M-Maha	0.64	0.64	0.00	99.90
	M-Maha-EF	0.63	0.63	0.00	99.90
Televisión por cable	U	0.99	0.98	0.01***	
	M-VC	0.99	0.99	0.00***	71.60
	M-Maha	0.99	0.99	0.00	100.00
	M-Maha-EF	0.99	0.99	0.00	100.00
Computador en la casa	U	0.67	0.52	0.15***	
	M-VC	0.70	0.71	-0.01**	93.30
	M-Maha	0.67	0.67	0.00	99.90
	M-Maha-EF	0.66	0.66	0.00	99.90
Vivir en una Casa	U	0.51	0.52	-0.01***	
	M-VC	0.55	0.57	-0.01**	-2.40
	M-Maha	0.52	0.52	0.00	99.50
	M-Maha-EF	0.51	0.51	0.00	99.30
Vivir en una Apt.	U	0.48	0.45	0.03***	
	M-VC	0.43	0.42	0.01**	60.60
	M-Maha	0.47	0.47	0.00	99.80
	M-Maha-EF	0.47	0.47	0.00	99.70
Casa propia pagada	U	0.41	0.38	0.03***	
	M-VC	0.43	0.43	0.00	95.60
	M-Maha	0.42	0.42	0.00	100.00
	M-Maha-EF	0.41	0.41	0.00	99.90
Casa propia en pago	U	0.41	0.44	-0.03***	
	M-VC	0.39	0.42	-0.02***	28.20
	M-Maha	0.41	0.41	0.00	100.00
	M-Maha-EF	0.41	0.41	0.00	99.80
Casa arrendada	U	0.10	0.06	0.03***	
	M-VC	0.10	0.09	0.01***	55.70
	M-Maha	0.09	0.09	0.00	100.00
	M-Maha-EF	0.09	0.09	0.00	99.90

La segunda columna muestra el tipo de emparejamiento utilizado. M-VC es vecino más cercano, M-Maha es usando la distancia de Mahalanobis y M-Maha-EF es usando el ultimo tipo de emparejamiento y agregando efectos fijos de tiempo y de municipio. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$\hat{\theta}_x Internet en el Trabajo$$

Cuadro 3: Variable dependiente logaritmo del ingreso laboral
Hipótesis 1: Cálculo para Colombia, comparables
con otros países de la región

	Internet en el Trabajo		
	Panel A - Hombres		
	(1)	(2)	(3)
	MCO	PSM	PSM
		Linea Base	Modelo Final
Asalariados	0.23*** (0.00)	0.21*** (0.01)	0.17*** (0.01)
Cuenta Propia	0.47*** (0.00)	0.43*** (0.03)	0.27*** (0.02)
	Panel B - Toda la muestra		
Asalariados	0.21*** (0.00)	0.16*** (0.01)	0.12*** (0.01)
Cuenta Propia	0.43*** (0.00)	0.46*** (0.02)	0.32*** (0.02)

La columna 1 muestra el modelo estimado con MCO (ecuación 1). La columna 2 muestra el modelo base de Navarro (2011) para Colombia, estimado con PSM (ecuación 4). La columna 3 muestra el modelo final estimado con PSM (ecuación 8). Error Estándar en paréntesis . * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$\hat{\theta}_x Lugar de Uso$$

Cuadro 4: Variable dependiente logaritmo del ingreso laboral. Estimado por **PSM**
Hipótesis 2: Heterogeneidades según lugar de uso del Internet

	Panel A - Hombres			
	(1)	(2)	(3)	(4)
	En el trabajo	En el hogar	En ambos lugares	Al menos un lugar
Asalariados	0.17*** (0.01)	0.10*** (0.01)	0.15*** (0.01)	0.17*** (0.01)
Cuenta Propia	0.27*** (0.02)	0.16*** (0.03)	0.30*** (0.02)	0.30*** (0.03)
	Panel B - Toda la muestra			
Asalariados	0.12*** (0.01)	0.10*** (0.01)	0.15*** (0.01)	0.16*** (0.01)
Cuenta Propia	0.32*** (0.02)	0.13*** (0.02)	0.32*** (0.02)	0.28*** (0.02)

Todos los modelos son estimados con el método del emparejamiento de Mahalanobis (ecuación 10). La columna 1 muestra el efecto del Internet en el Trabajo. La columna 2 muestra el efecto del Internet en el Hogar. La columna 3 muestra el efecto del uso en ambos lugares. La columna 4 muestra el efecto de al menos un lugar. Error Estándar en paréntesis . * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$\hat{\theta}_x \text{ManejodeInternet}$$

Cuadro 5: Variable dependiente logaritmo del ingreso laboral. Estimado por **PSM**.

Hipótesis 3: Manejo del Internet e intensidad del uso

Panel A - Hombres				
	(1)	(2)	(3)	(4)
	Educación	Entretenimiento	Educación e intensidad diaria	Entretenimiento e intensidad diaria
Asalariados	0.03 (0.02)	0.02 (0.02)	0.04*** (0.01)	0.01 (0.02)
Cuenta Propia	0.06 (0.05)	0.02 (0.01)	0.07*** (0.02)	0.01 (0.01)
Panel B - Toda la muestra				
Asalariados	0.02 (0.02)	0.01 (0.01)	0.04*** (0.01)	0.00 (0.01)
Cuenta Propia	0.05 (0.04)	0.01 (0.02)	0.06*** (0.02)	0.01 (0.02)

La columna 1 muestra el efecto del manejo del Internet en educación. La columna 2 muestra el efecto del manejo del Internet en entretenimiento (ecuación 11). La columna 3 muestra la interacción con manejo diario en Educación. La columna 4 muestra la interacción con manejo diario en Entretenimiento (Ecuación 12). Error Estándar en paréntesis. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$\hat{\theta}_x \text{InternetenelTrabajo} * \text{NivelEducativo}$$

Cuadro 6: Variable dependiente logaritmo del ingreso laboral. Estimado por **PSM**.

Hipótesis 4: Efecto de la interacción de Internet en el trabajo y niveles educativos

Panel A - Hombres				
	(1)	(2)	(3)	(4)
	Secundaria Completa	Universidad Incompleta	Universidad Completa	Posgrado
Asalariados	0.12*** (0.02)	0.17*** (0.03)	0.35*** (0.02)	0.57*** (0.02)
Cuenta Propia	0.17*** (0.03)	0.20*** (0.03)	0.29*** (0.03)	0.47*** (0.03)
Panel B - Toda la muestra				
Asalariados	0.11*** (0.02)	0.17*** (0.03)	0.32*** (0.02)	0.54*** (0.02)
Cuenta Propia	0.18*** (0.03)	0.22*** (0.03)	0.33*** (0.02)	0.48*** (0.03)

El grupo de control en cada tratamiento es las personas con secundaria incompleta y sin Internet. La columna 1 es tener bachillerato completo. La columna 2 es tener universidad incompleto. La columna 3 es tener universidad completa. La columna 4 es tener posgrado (ecuación 13). Error Estándar en paréntesis . * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

8. Apéndice

Figura A1: Actividad Económica

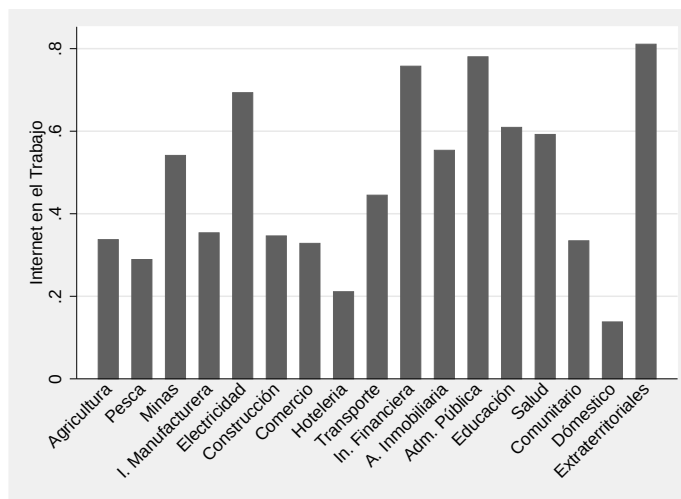


Figura A2: Ocupación

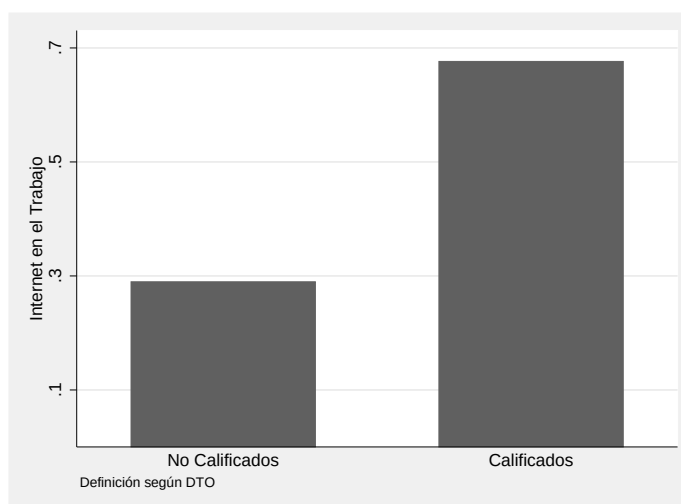
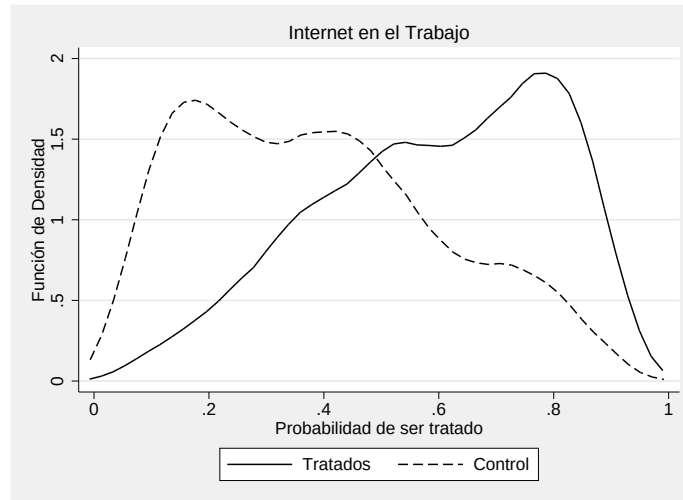
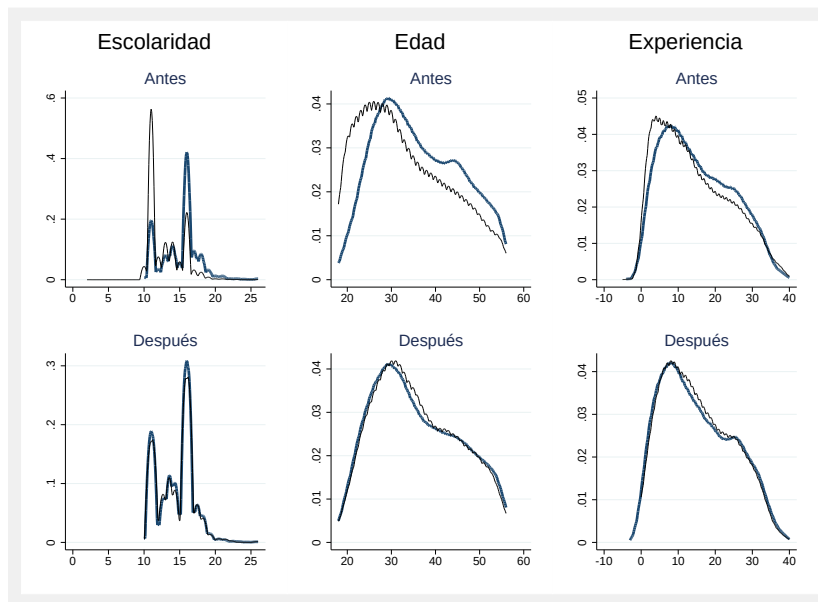


Figura A3: Propensity Score para Empleados Cuenta Propia



PS con el modelo final (Ecuación 7). El soporte común incluye el 91 % de los datos de la muestra. Las observaciones que no se incluyen pertenecen a el grupo de tratamiento.

Figura A4: Propensity Score para Empleados Cuenta Propia



Estimaciones usando el PSM (Ecuación 7 y 8). Línea Oscura grupo de Tratamiento. Línea clara grupo de control.

Cuadro A1: Diferencia de medias entre cuenta propia y asalariado

Variable	Cuenta Propia	Asalariado	Diferencia
Logaritmo del Ingreso Laboral	13.29 (0.97)	13.69 (0.68)	-0.40*** (0.00)
Internet en Trabajo	0.28 (0.45)	0.59 (0.49)	-0.31*** (0.00)
Internet en el Hogar	0.51 (0.5)	0.51 (0.5)	-0.01*** (0.00)
Internet en ambas partes	0.16 (0.36)	0.32 (0.46)	-0.16*** (0.00)
Internet en al menos una de las partes	0.62 (0.48)	0.78 (0.41)	-0.16*** (0.00)
Doce o más años de educación	0.61 (0.49)	0.66 (0.47)	-0.05*** (0.00)
Edad	33.54 (9.14)	33.35 (9.44)	0.19*** (0.04)
Edad al cuadrado	1208.13 (660.24)	1201.29 (688.59)	6.83** (2.98)
Teléfono Fijo	0.57 (0.5)	0.61 (0.49)	-0.05*** (0.00)
Televisión por cable	0.98 (0.14)	0.99 (0.11)	-0.01*** (0.00)
Computador en la casa	0.59 (0.49)	0.61 (0.49)	-0.02*** (0.00)
Vivir en una Casa	0.56 (0.5)	0.52 (0.5)	0.04*** (0.00)
Vivir en una Apt.	0.42 (0.49)	0.46 (0.5)	-0.04*** (0.00)
Casa propia pagada	0.42 (0.49)	0.41 (0.49)	0.01*** (0.00)
Casa arrendada	0.4 (0.49)	0.41 (0.49)	-0.01*** (0.00)
Casa propia en pago	0.06 (0.24)	0.09 (0.28)	-0.02*** (0.00)
Casado	0.27 (0.44)	0.29 (0.46)	-0.02*** (0.00)
Salud Subsidiada	0.83 (0.38)	0.96 (0.2)	-0.13*** (0.00)
Pago de Pensión	0.25 (0.43)	0.82 (0.38)	-0.57*** (0.00)
Actividad Económica	10.55 (3.14)	10.78 (3.29)	-0.23*** (0.02)
Ocupación	43.02 (28.93)	38.37 (24.45)	4.64*** (0.11)

Errores estándar en paréntesis. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

Cuadro A2: Diferencia de medias entre tratados y no tratados antes y después del emparejamiento. Modelo para empleados Cuenta Propia

Variable	Antes (U)	Promedio		Diferencia	%Reducción del Sesgo
	Después (M)	Tratados	No Tratados		
Doce o más años de educación	U	0.83	0.51	0.32***	
	M-VC	0.86	0.87	-0.01	96.20
	M-Maha	0.82	0.82	0.00	100.00
	M-Maha-EF	0.82	0.82	0.00	100.00
Edad	U	35.26	32.86	2.40***	
	M-VC	35.02	34.49	0.52***	78.20
	M-Maha	35.30	35.29	0.01	99.50
	M-Maha-EF	35.46	35.45	0.01	99.60
Edad al cuadrado	U	1327	1169	157.60***	
	M-VC	1309	1274	35.20**	77.70
	M-Maha	1332	1331	1.30	99.20
	M-Maha-EF	1344	1343	1.10	99.30
Teléfono Fijo	U	0.62	0.52	0.10***	
	M-VC	0.63	0.63	0.00	98.80
	M-Maha	0.63	0.63	0.00	99.90
	M-Maha-EF	0.63	0.63	0.00	99.90
Televisión por cable	U	0.99	0.97	0.01***	
	M-VC	0.99	0.98	0.01**	55.30
	M-Maha	0.99	0.99	0.00	100.00
	M-Maha-EF	0.99	0.99	0.00	100.00
Computador en la casa	U	0.69	0.52	0.17***	
	M-VC	0.72	0.70	0.02**	90.00
	M-Maha	0.70	0.70	0.00	99.90
	M-Maha-EF	0.68	0.68	0.00	99.90
Vivir en una Casa	U	0.53	0.58	-0.05***	
	M-VC	0.57	0.60	-0.03**	42.90
	M-Maha	0.54	0.54	0.00	99.80
	M-Maha-EF	0.54	0.54	0.00	99.80
Vivir en una Apt.	U	0.46	0.40	0.06***	
	M-VC	0.42	0.39	0.03**	56.50
	M-Maha	0.45	0.45	0.00	99.80
	M-Maha-EF	0.45	0.45	0.00	99.80
Casa propia pagada	U	0.43	0.42	0.01***	
	M-VC	0.46	0.50	-0.04***	-375.70
	M-Maha	0.44	0.44	0.00	100.00
	M-Maha-EF	0.44	0.44	0.00	100.00
Casa propia en pago	U	0.41	0.40	0.01***	
	M-VC	0.38	0.35	0.03***	-274.50
	M-Maha	0.39	0.39	0.00	100.00
	M-Maha-EF	0.39	0.39	0.00	100.00
Casa arrendada	U	0.07	0.05	0.02***	
	M-VC	0.08	0.06	0.02***	13.70
	M-Maha	0.08	0.08	0.00	100.00
	M-Maha-EF	0.07	0.07	0.00	100.00

La segunda columna muestra el tipo de emparejamiento utilizado. M-VC es vecino más cercano, M-Maha es usando la distancia de Mahalanobis y M-Maha-EF es usando el ultimo tipo de emparejamiento y agregando efectos fijos de tiempo y de municipio. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$\hat{\theta}_x \text{Internet en el Trabajo}$$

Cuadro A3: Variable dependiente logaritmo del ingreso laboral

Eliminado soporte común				
Internet en el Trabajo				
Panel A - Hombres				
	(1)	(2)	(3)	(4)
	PSM	PSM	PSM	PSM
	Modelo Final	5 %	10 %	20 %
Asalariados	0.17*** (0.00)	0.17*** (0.00)	0.18*** (0.01)	0.20*** (0.01)
Cuenta Propia	0.27*** (0.00)	0.27*** (0.00)	0.27*** (0.03)	0.28*** (0.02)
Panel B - Toda la muestra				
Asalariados	0.12*** (0.00)	0.12*** (0.00)	0.14*** (0.01)	0.15*** (0.01)
Cuenta Propia	0.32*** (0.00)	0.32*** (0.00)	0.33*** (0.02)	0.34*** (0.02)

Estimaciones con ecuación 8. La columna 1 muestra el modelo final con todo el soporte común. La columna 2 muestra el modelo eliminando el 5 % del soporte común. La columna 3 muestra el modelo eliminando el 10 % del soporte común. La columna 4 muestra el modelo eliminando el 20 % del soporte común. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$H_0: \hat{\theta}_x \text{Amboslugares} - \hat{\theta}_x \text{Al menos un lugar} = 0$$

Cuadro A4: Test de diferencia de modelos entre uso del Internet en el trabajo y en el hogar comparado con al menos un solo lugar

	En ambos lugares	Al menos un lugar	Diferencia de medias	Diferencia de Modelos
	(1)	(2)	(3)	(4)
Panel A - Hombres				
Asalariado	0.15 (0.01)	0.17 (0.01)	-0.02* (0.02)	-0.02** (0.05)
Cuenta Propia	0.30 (0.02)	0.30 (0.03)	0.00 (0.04)	0.00 (0.18)
Panel B - Toda la muestra				
Asalariado	0.15 (0.01)	0.16 (0.01)	-0.01 (0.02)	-0.01* (0.01)
Cuenta Propia	0.32 (0.02)	0.28 (0.02)	0.04** (0.04)	0.04** (0.03)

La columna 1 y 2 muestran los coeficientes de estimar la ecuación 10 para usar Internet en ambos lugares y en al menos uno. La columna 3 expone la prueba de diferencia de medias, que se distribuye como una $t - student$ y se hace la prueba a dos colas. La columna 4 muestra la prueba de diferencia de modelos se basa en un test de Wald, que se distribuye como una Chi^2 . Error Estándar en paréntesis. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$\hat{\theta}_x InternetenEducacion * FrecuanciadeUso$$

Cuadro A5: Variable dependiente logaritmo del ingreso laboral. Estimado por **PSM**.

Hipótesis 3: Manejo del Internet para educación e intensidad del uso

	Educación e interacción con frecuencia			
	Diaria (1)	Semanal (2)	Mensual (3)	Menor que mensual (4)
Panel A - Hombres				
Asalariados	0.04*** (0.01)	0.02*** (0.01)	0.01 (0.01)	0.00 (0.01)
Cuenta Propia	0.07*** (0.01)	0.03*** (0.01)	0.02 (0.02)	0.01 (0.01)
Panel B - Toda la muestra				
Asalariados	0.04*** (0.01)	0.02*** (0.01)	0.01 (0.01)	0.00 (0.01)
Cuenta Propia	0.06*** (0.02)	0.03*** (0.01)	0.01 (0.02)	0.01 (0.01)

Los resultados son estimados con la ecuación 12. La columna 1 muestra el impacto la interacción entre educación e Internet diariamente. La columna 2 la interacción con Internet semanalmente, pero no diariamente. La columna 3 la interacción con Internet mensualmente, pero no semanalmente. La columna 4 la interacción con Internet menor a mensualmente. Error Estándar en paréntesis. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$\hat{\theta}_x LugardeUso$$

Cuadro A6: Variable dependiente logaritmo del ingreso laboral. Estimado por MCO y PSM.

Hipótesis 2: Manejo del Internet e intensidad del uso

	Panel A - Hombres							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	En el trabajo MCO	En el trabajo PSM	En el hogar MCO	En el hogar PSM	En ambos lugares MCO	En ambos lugares PSM	Al menos un lugar MCO	Al menos un lugar PSM
Asalariados	0.20*** (0.02)	0.17*** (0.01)	0.12*** (0.02)	0.10*** (0.01)	0.16*** (0.01)	0.15*** (0.01)	0.19*** (0.02)	0.17*** (0.01)
Cuenta Propia	0.31*** (0.04)	0.27*** (0.02)	0.17*** (0.02)	0.16*** (0.03)	0.30*** (0.02)	0.30*** (0.02)	0.34*** (0.04)	0.30*** (0.03)
Panel B - Toda la muestra								
Asalariados	0.16*** (0.01)	0.12*** (0.01)	0.14*** (0.02)	0.10*** (0.01)	0.16*** (0.01)	0.15*** (0.01)	0.18*** (0.02)	0.16*** (0.01)
Cuenta Propia	0.39*** (0.03)	0.32*** (0.02)	0.15*** (0.02)	0.13*** (0.02)	0.32*** (0.02)	0.32*** (0.02)	0.30*** (0.03)	0.28*** (0.02)

Estimaciones con la Ecuación 10. Error Estándar en paréntesis. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$\hat{\theta}_x \text{ManejodelInternet}$$

Cuadro A7: Variable dependiente logaritmo del ingreso laboral. Estimado por MCO y PSM.

Hipótesis 3: Manejo del Internet e intensidad del uso								
Panel A - Hombres								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Educación		Entretenimiento		Educación e intensidad diaria		Entretenimiento e intensidad diaria	
	MCO	PSM	MCO	PSM	MCO	PSM	MCO	PSM
Asalariados	0.04**	0.03	0.03**	0.02	0.09***	0.04***	0.01*	0.01
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
Cuenta Propia	0.07**	0.06	0.02**	0.02	0.11***	0.07***	0.01*	0.01
	(0.05)	(0.05)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Panel B - Toda la muestra								
Asalariados	0.03*	0.02	0.02**	0.01	0.10***	0.04***	0.00*	0.00
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Cuenta Propia	0.06**	0.05	0.03**	0.01	0.12***	0.06***	0.01*	0.01
	(0.04)	(0.04)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)

Estimaciones con la Ecuación 11 y 12. Error Estándar en paréntesis. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

$$\hat{\theta}_x \text{InternetenelTrabajo} * \text{NivelEducativo}$$

Cuadro A8: Variable dependiente logaritmo del ingreso laboral. Estimado por MCO y PSM.

Hipótesis 4: Efecto de la interacción de Internet en el trabajo y niveles educativos

Panel A - Hombres								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Secundaria completa		Universidad incompleta		Universidad completa		Posgrado	
	MCO	PSM	MCO	PSM	MCO	PSM	MCO	PSM
Asalariados	0.18***	0.12***	0.24***	0.17***	0.39***	0.35***	0.62***	0.57***
	(0.01)	(0.02)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)
Cuenta Propia	0.21***	0.17***	0.26***	0.20***	0.35***	0.29***	0.51***	0.47***
	(0.01)	(0.03)	(0.02)	(0.03)	(0.01)	(0.03)	(0.02)	(0.03)
Panel B - Toda la muestra								
Asalariados	0.16***	0.11***	0.23***	0.17***	0.38***	0.32***	0.66***	0.54***
	(0.01)	-0.02	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.02)
Cuenta Propia	0.21***	0.18***	0.25***	0.22***	0.35***	0.33***	0.51***	0.48***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.01)	(0.03)

Estimaciones con la Ecuación 13. Error Estándar en paréntesis. * representa significancia estadística al 10 %, ** al 5 % y *** al 1 %.

9. Referencias

1. Acemoglu, D. 1998. Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality. *Quarterly Journal of Economics*.
2. Angrist, J and Pischke, J. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
3. Arellano, M. and Bond, S. 1991. Some Tests of Specification for Panel Data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*.
4. Autor, D; Lawrence, F. and Krueger, A. 1998. Computing Inequality: Have Computers Changed the Labor Market? *Quarterly Journal of Economics*.
5. Autor, D; Levy, F and Murnane, R. 2003. The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics*.
6. Autor, D; Lawrence F. and Kearney, M. 2006. The Polarization of the U.S. Labor Market. *American Economic Review*.
7. Bartel, A; Ichniowski, C. and Shaw, K. 2007. How Does Information Technology Affect Productivity? Plant-Level Comparisons of ProductInnovation, Process Improvement, and Worker Skills. *The Quaterly Journal of Economics*.
8. Benavente, J; Bravo, D. and Montero, R. 2011. Wages and workplace computer use in Chile. *The Developing Economies*.
9. Benavides, J; Castro, F; Devis, L; Olivera, M. 2011. Impacto de las Tecnologías de la Información y las Comunicaciones (TIC) en el Desarrollo y la Competitividad del País. Fedesarrollo.
10. Bernal, R and Peña, X. 2012. *Guía Practica para la Evaluación de Impacto*. Universidad del los Andes. Facultad de Economía. Centro de Estudios sobre Desarrollo Económico (CEDE).
11. Blundell, R; Dearden, L. and Sianesi, B. 2005. Evaluating the Effect of Education on Earnings: Models, Methods and Results from the National Child Development Survey. *Journal of the Royal Statistical Society*.
12. Bureau of labor statistics (BLS). 2010. *Standard Occupational Classification (SOC)*.
13. Chu, S. 2013. Internet, Economic Growth and Recession. *Modern Economy*.

14. Departamento Administrativo Nacional de Estadística (DANE). 2011. Gran Encuesta Integrada de Hogares (GEIH). Archivo Nacional de Datos.
15. Dearden, L; Machin, S. and Vignoles, A. 2009. Economics of Education Research: A Review and Future Prospects. Oxford Review of Education.
16. Dehejia, R. and Wahba, S. 2002. Propensity Score-Matching Methods for Nonexperimental Causal Studies. The Review of Economics and Statistics.
17. DiMaggio, P. and Bonikowski, B. 2008. Make Money Surfing the Web? The Impact of Internet Use on the Earnings. American Sociological Review.
18. Forman, C; Goldfarb, A. and Greenstein, S. 2012. The Internet and Local Wages: A Puzzle. American Economic Review
19. Fountain, J. 2003. The Virtual State: Transforming America government? National Civil Review.
20. Freeman, R. 2002, The Labour Market in the New Information Economy. Oxford Review of Economic Policy.
21. Goss, E. and Phillips, E. 2002. How Information Technology Affects Wages: Evidence Using Internet Usage As a Proxy for IT Skills. Journal of Labour Research
22. Gutiérrez, L. 2011. ICT and labor productivity in Colombian manufacturing industry. Chapter 1. Part B. ICT In Latin America. A microdata analysis. United Nations.
23. Heckman, J; Ichimura, H and Todd, P. 1998. Matching as an Econometric Evaluation Estimator. Review of Economic Studies.
24. Heckman, J; LaLonde, R and Smith, J. 1999. The Economics and Econometrics of Active Labour Market Programs. Handbook of Labour Economics.
25. Hotz, J. and Miller, R. 1988. An Empirical Analysis of Life Cycle Fertility and Female Labor Supply. Econometrica.
26. Katz, R. and Callorda, F. 2011. Medicion de Impacto del Plan Vive Digital en Colombia y de la Masificacion de Internet en la Estrategia de Gobierno en Linea. Cintel.
27. Katz, R and Koutroumpis, P. 2012. The Economic Impact of Broadband in Panama. Telecommunication Developmente Sector.
28. Katz, R and Callorda, F. 2013. Economic Impact of Broadband Deployment in Ecuador. IRSI.

29. Krueger, A. 1993. How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989. *Quarterly Journal of Economics*.
30. Koutroumpis, P. 2009. The Economic Impact of Broadband on Growth: A Simultaneous Approach. *Telecommunications Policy*
31. Link, C. and Ratledge, E. 1975. The Influence of the Quantity and Quality of Education on Black-White Earnings Differentials: Some New Evidence. *The Review of Economics and Statistics*.
32. Mincer, J. 1958. Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*.
33. Ministerio de Tecnologías de la Información y las Comunicaciones. 2011. Boletín Trimestral de la TIC.
34. Navarro, L. 2011. Impact of Internet use on individual earnings. Chapter 3. Part A. *ICT In Latin America. A microdata analysis*. United Nations.
35. Prasad, V and Ramamurthy, K. 2001. The Influence of Internet – Marketing Integration on Marketing Competencies and Export Performance. *Journal of International Marketing*.
36. Qiang, C. and Rossotto, C. 2009. Economic Impacts of Broadband. *American Sociological Review*.
37. Rosenbaum, P and Rubin, D. 1983. The Central Role of Propensity Score in Observational Studies of Causal Effect. *Biometrika*.
38. Sianesi, B. and Leuven, E. (2003). psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. <http://ideas.repec.org/c/boc/bocode/s432001.html>., version.
39. Tilley, P. 2005. Gender difference in deception and its detection under varying electronic media conditions. *System Sciences*.
40. Zahra, S; Rawhouse, H and Bhawe, N. 2008. Globalization of social entrepreneurship. *Strategic Entrepreneurship Journal opportunities*.